

A TOOLKIT FOR

# Centering Racial Equity Throughout Data Integration



Societal “progress” is often marked by the construction of new infrastructure that fuels change and innovation. Just as railroads and interstate highways were the defining infrastructure projects of the 1800 and 1900s, the development of data infrastructure is a critical innovation of our century.

Railroads and highways were drivers of development and prosperity for some investors and sites. Yet other individuals and communities were harmed, displaced, bypassed, ignored, and forgotten by those efforts.

**As railroads and highways both developed and decimated communities, so too can data infrastructure.**

At this moment in our history, we can co-create data infrastructure to promote racial equity and the public good, or we can invest in data infrastructure that disregards the historical, social, and political context—reinforcing racial inequity that continues to harm communities. Building data infrastructure without a racial equity lens and understanding of historical context will exacerbate existing inequalities along the lines of race, gender, class, and ability. Instead, we commit to contextualize our work in the historical and structural oppression that shapes it, and organize stakeholders across geography, sector, and experience to center racial equity throughout data integration.

# Introduction

Data sharing and data integration to inform decision making across government entities is now commonplace, and occurs at every level—local, state, and federal. While most data sharing and integration occurs within a legal and governance framework, an emphasis on racial equity, transparency, and community engagement is often peripheral. This is especially troubling because government policies and programs that produce administrative data have often played a direct role in creating, enabling, and sustaining institutional and structural racism.

We aim to change this.

With trust in government and experts (e.g., researchers) at historic lows,<sup>1</sup> efforts that rely exclusively on these institutions to “use data to solve social problems” are unlikely to succeed. Too often government organizations and their research partners fail to identify and address issues of bias in data. Further, even if such issues are identified, these organizations are not equipped to repair trust with communities that have experienced harm.

This body of work seeks to encourage shifts of awareness and practice, by centering racial equity and community voice within the context of data integration and use. Our vision is one of ethical data use with a racial equity lens, that supports power sharing and building across agencies and community members.

- **Racial equity** is the condition where one’s racial identity no longer influences how one fares in society. This includes the creation of racially just policies, practices, attitudes, and cultural messages, and the elimination of structures that reinforce differential experiences and outcomes by race.<sup>2</sup>
- **Administrative data** are data collected during the routine process of administering programs.
- **Administrative data reuse** involves using these data in a way not originally intended (e.g., for research).
- **Cross-sector data sharing** is the practice of providing access to information not otherwise available across agencies.
- **Data integration** involves data sharing that includes identifiable information (e.g, name, date of birth, SSN), so that records can be linked, or integrated at the individual level.

Data sharing and data integration involve significant privacy risks, and all data use should be carefully considered to ensure sharing is legal and ethical, with a purpose that can be linked to action to improve outcomes.

1 Rainie, L., Keeter, S., & Perrin, A. (2019, December 31)

2 Racial Equity Tools (n.d.), Glossary

## Why Data Infrastructure + Racial Equity?

Cross-sector data sharing and integration enable the transformation of individual-level information into actionable intelligence that can be used to understand urgent and long-term community needs; improve services, systems, and practices; develop innovative policies and interventions; and, ultimately, build stronger communities. **Yet, the way that cross-sector data are used can also reinforce legacies of racist policies and produce inequitable resource allocation, access, and outcomes.**

We understand structural racism as the normalization and legitimization of historical, cultural, institutional, and interpersonal dynamics that advantage Whites, “while producing cumulative and chronic adverse outcomes for People of Color.” Embedded within structural racism is institutional racism, the ways “policies and practices of organizations or parts of systems (schools, courts, transportation, etc.) create different outcomes for different racial groups”<sup>3</sup> (see Appendix II).

Black, Indigenous, and people of color (BIPOC)<sup>4</sup> and/or people living in poverty are often over-represented within government agency data systems, and disparate representation in data can cause disparate impact.<sup>5</sup> Laws, policies, business rules, and narratives are permeated by structural racism, which is the root cause of the racial disparities evident in system outcomes. Such disparities are often sterilized by well-intentioned names (e.g., “disproportionate contact” in criminal justice or the “achievement gap” in education) that hide the social consequence of structural racism: that, as a group, Black, Indigenous, and people of color in America have worse outcomes in many human service system outcome measures regardless of socioeconomic status.<sup>6</sup> And yet, many agency solutions and data initiatives are largely disconnected from this root cause, and the “hunt for more data is [often] a barrier for acting on what we already know.”<sup>7</sup>

<sup>3</sup> Racial Equity Tools (n.d.), Core Concepts: Racism

<sup>4</sup> We intentionally use the acronym BIPOC (Black, Indigenous, people of color) as a term that seeks to recognize the unique experience of Black and Indigenous People within the United States. We recognize that naming is power, and we remain committed to using language that supports pro-Blackness and Native visibility, while dismantling white supremacy.

<sup>5</sup> Barocas, S., & Selbst, A. D. (2016)

<sup>6</sup> Hayes-Greene, D., & Love, B. P. (2018)

<sup>7</sup> Benjamin, R. (2019)

With this knowledge, we call for re-users of administrative data to center racial equity in data practice. We call for the inclusion of community voices and power sharing at every stage of design, use, and implementation. We call for relationship building among those represented in the data and those using the data. Without a deliberate effort to address structural racism, institutional racism, and unrecognized bias, data integration will inevitably reproduce and exacerbate existing harm.

**To avoid this, we must embed questions of racial equity throughout the data life cycle:**

- In planning
- In data collection
- In data access
- In algorithms/use of statistical tools
- In data analysis
- In reporting and dissemination

We are at a pivotal moment, one in which the use of data is accelerating in both exciting and concerning ways. We have access to greater amounts of data than at any other point in our history, but privacy laws and practice lag behind, placing Black, Indigenous, and communities of color at the greatest risk of the “data-ification of injustice.”<sup>8</sup>

Acknowledging history, harm, and the potentially negative implications of data integration for groups marginalized by inequitable systems is a key first step, but it is only a first step. To go beyond this, we must center the voices, stories, expertise, and knowledge of these communities in decision making, and take collective action with shared power to improve outcomes and harness data for social good.

We are working to create a new kind of data infrastructure—one that dismantles ‘feedback loops of injustice’<sup>9</sup> and instead shares power and knowledge with those who need systems change the most.

Will you join us?

<sup>8</sup> Benjamin, R. (2019)

<sup>9</sup> Eubanks, V. (2018)

# Contents

1	<b>Introduction</b>
2	Why Data Infrastructure + Racial Equity?
5	Who We Are
7	<b>How to Use This Toolkit</b>
7	Who Should Use This Toolkit
8	Get Started
12	<b>Benefits, Limitations, &amp; Risks</b>
14	<b>Centering Racial Equity Throughout the Data Life Cycle</b>
15	Racial Equity in Planning
18	Racial Equity in Data Collection
21	Racial Equity in Data Access
25	Racial Equity in Algorithms/Statistical Tools
28	Racial Equity in Data Analysis
30	Racial Equity in Reporting and Dissemination
33	What Next?
34	<b>Toolkit Activity 1: Who Should Be at the Table?</b>
40	<b>Toolkit Activity 2: Mapping Assets and Engaging Community</b>
44	<b>Toolkit Activity 3: Identifying Root Causes through Factor Analysis</b>
47	<b>References</b>
50	<b>Appendix I: Work in Action Throughout the Data Life Cycle</b>
50	Racial Equity in Planning: Work in Action
53	Racial Equity in Data Collection: Work in Action
56	Racial Equity in Data Access: Work in Action
59	Racial Equity in Algorithms / Use of Statistical Tools: Work in Action
61	Racial Equity in Data Analysis: Work in Action
63	Racial Equity in Reporting and Dissemination: Work in Action
68	<b>Appendix II: Terms</b>
70	<b>Appendix III: Expanded Civic Data User Profiles</b>
72	<b>Appendix IV: Resources</b>

## Who We Are

We are a diverse group of civic data stakeholders, including community advocates, staff of local and national nonprofit organizations, university-based and applied researchers, state and local government administrators and analysts, foundation staff, and service providers. We have worked together since May 2019 to co-create strategies and identify best practices for administrative data reuse for government agencies, nonprofit organizations, and data collaboratives that center racial equity and gives power to community voice. Specifically, our work seeks to help agencies acknowledge and compensate for the harms and bias baked into data and data structures, into practice, and into cultural understandings and perceptions of populations served by government agencies.

### WORKGROUP CONTRIBUTORS

Niiobli Armah, My Brother's Keeper, Equity Intelligence Platform  
 Bridget Blount, Baltimore's Promise  
 Angela Bluhm, Chief Education Office, State of Oregon  
 Katy Collins, Allegheny County Department of Human Services  
 Sheila Dugan, GovEx, Johns Hopkins University  
 Sue Gallagher, Broward Data Collaborative, Children's Services Council of Broward County  
 Laura Jones, Writer and Community Advocate based in Minnesota  
 Chris Kingsley, Annie E. Casey Foundation  
 Ritika Sharma Kurup, StriveTogether  
 Tamika Lewis, Our Data Bodies  
 Rick Little, Utah Dept of Human Services, Management Information Center  
 Tawana Petty, Detroit Community Technology Project & Our Data Bodies  
 Raintry Salk, Race Forward and Government Alliance for Racial Equity (GARE)  
 Michelle Shevin, Ford Foundation

### SITE-BASED CONTRIBUTORS

Allegheny County (PA), Department of Human Services, Office of Analytics, Technology, & Planning,  
 Samantha Loaney, Brian Bell, Ellen Kitzerow, Julia Reuben, Shannon Flynn, & Jamaal Davis  
 Allegheny County (PA) Department of Human Services, Office of Equity & Inclusion,  
 Shauna Lucadamo & Jessica Ruffin  
 Automating.NYC, Deepra Yusuf, Elyse Voegeli, Akina Younge, & Jon Truong  
 Birth through Eight Strategy for Tulsa (BEST), Jessica England & Dan Sterba  
 Children's Services Council of Broward County (FL), Sue Gallagher  
 City of Asheville (NC), Christen McNamara & Kimberlee Archie  
 City of Tacoma (WA), Alison Beason  
 DataWorks NC, Libby McClure & John Killeen

Kentucky Center for Statistics, Jessica Cunningham  
Mecklenburg County (NC) Community Support Services, Courtney LaCaria & Mary Ann Priester  
New York City Administration for Children's Services & Youth Studies Programs at the CUNY School  
of Professional Studies, Sarah Zeller-Berkman  
Take Control Initiative (OK), Emma Swepston, Laura Bellis, & Brandy Hammons

## AISP CONTRIBUTORS

Lead author: Amy Hawn Nelson, PhD  
Della Jenkins  
Sharon Zanti  
Matthew Katz  
Emily Berkowitz  
TC Burnett  
Dennis Culhane, PhD

We also recognize individuals who provided external review of this toolkit, including Jessica Cunningham, Marcus Gaddy, Logan Koepke, Michelle Massie, and Jasmine McNealy.

## ACKNOWLEDGEMENTS

The authors wish to express their gratitude to the Annie E. Casey Foundation, as well as the other members of the Data Funders Collaborative who supported this work.

### Suggested Citation:

Hawn Nelson, A., Jenkins, D., Zanti, S., Katz, M., Berkowitz, E., et al. (2020). [A Toolkit for Centering Racial Equity Throughout Data Integration](#). Actionable Intelligence for Social Policy, University of Pennsylvania.



# How to Use This Toolkit

This toolkit is designed to help guide partnerships, collaboratives, agencies, and community initiatives seeking to center racial equity while using, sharing, and integrating administrative data. Not sure what we mean by using, sharing, and integrating administrative data? Take some time to review our Introduction to Introduction to [Data Sharing & Integration](#),<sup>10</sup> which covers key terms, concepts, and first steps.

## Who Should Use This Toolkit

We believe that all voices are needed in conversations about racial equity and data use, and the information presented here can be used by anyone. This toolkit and activities are specifically structured to support users of administrative data for civic purposes (see Appendix III) in their efforts to center racial equity. **These users could include:**

- **Members of institutions:** university-based researchers; government-agency administrators and analysts; foundation staff
- **Community advocates and members:** community and religious leaders; civic and neighborhood association members; students and parents/caregivers
- **Bridges between community and organizations:** service providers, social workers, case workers; staff of local backbone organizations; independent applied researchers

**Many types of civic data use are relevant to this toolkit, including:**

- Open data (data that can be shared without legal agreements in place)
- Protected administrative data (confidential data that can be shared only with sufficient security provisions in place, including data sharing agreements)
- Dashboards (administrative data aggregated to topic/indicator/subgroup/population)
- Integrated data systems (systems that regularly link administrative data across government agencies to improve programs and practices through evidence-based collaboration)
- Neighborhood indicators (data aggregated to place)
- Research, evaluation, and outcome measurement using administrative data
- Tools created by using administrative data, such as risk indicator scores

<sup>10</sup> See [Introduction to Data Sharing & Integration](#). (2020). Actionable Intelligence for Social Policy.

**WE STRONGLY ENCOURAGE:**

- Inclusive participatory governance around data access and use
- Social license for data access and use
- A developmental approach to data sharing and integration—start small and grow

**WE DISCOURAGE:**

- Broad access to individual-level linked data
- Data use for enforcement or investigation actions against residents
- Use of predictive algorithms without determining responsibility, explainability, accuracy, auditability, and fairness<sup>11</sup>
- Use of linked data across institutions that have patterns of institutional racism, specifically, law enforcement, which has demonstrated significant racialized harm without sufficient safeguards in place

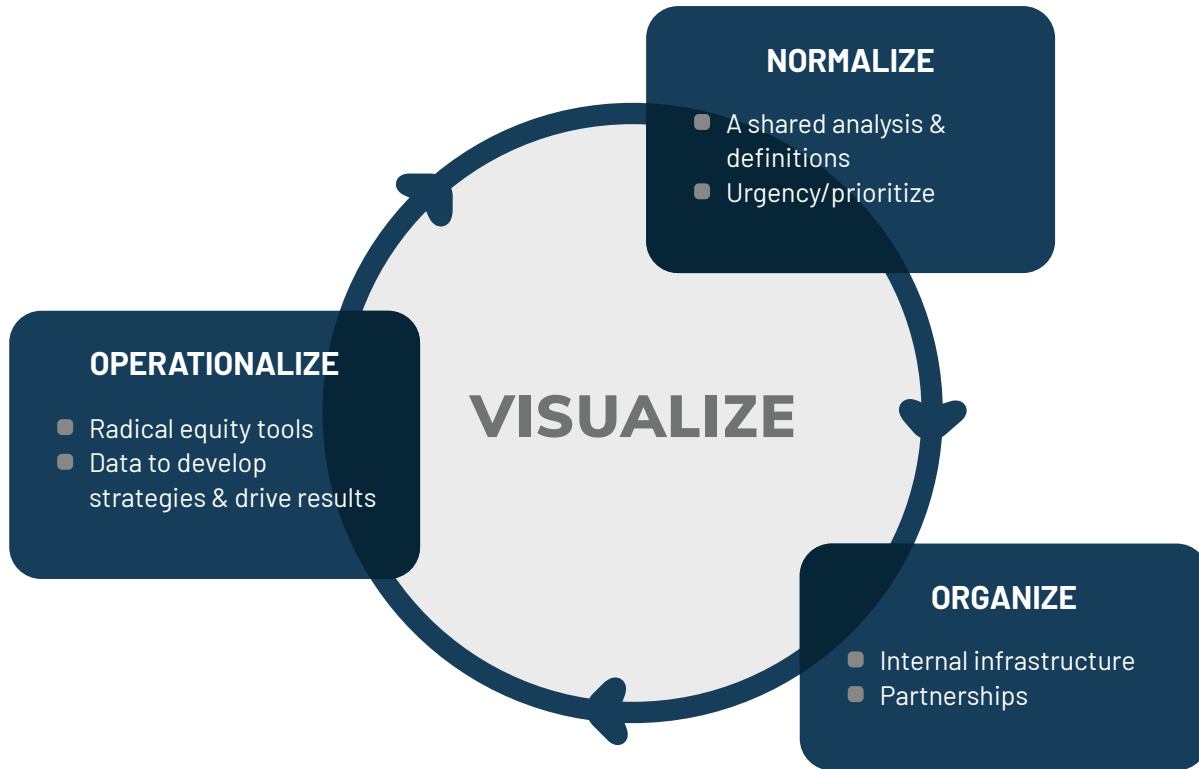
## Get Started

We recommend use of the following framework outlined by the Government Alliance for Racial Equity (GARE) to **normalize, organize, and operationalize**<sup>12</sup> racial equity throughout data integration.

We urge you to begin by thinking through the questions listed below to better understand your individual and institutional starting point for centering racial equity throughout data integration. Next, use the core questions in each of the Normalize, Organize, and Operationalize sections to reflect on ways that you and your organization can continue to grow in these areas. The links and activities included in each section will help you further assess and take action toward centering racial equity. It is important to note that this work is not linear, but iterative. This toolkit is intentionally modular, and we encourage you to use sections as needed in order to move the work of your organization forward.

<sup>11</sup> Diakopoulos, N., et al. (2017)

<sup>12</sup> Nelson, J., & Brooks, L. (2015)



Source: [GARE Communications Guide](#), May 2018.

## Normalize

**How has your lead agency/collaborative acknowledged the importance of a racial equity lens and demonstrated a commitment to engage in data integration efforts that are legal, are ethical, and center equity?**

**Assess your current activities using these resources:**

- [Data Ethics Workbook](#), UK Department for Digital, Culture, Media & Sport
- [Tool for Organizational Self-Assessment Related to Racial Equity](#), Coalition of Communities of Color

**We recommend these resources to move forward:**

- [Advancing Racial Equity and Transforming Government: A Resource Guide to Put Ideas into Action](#), Government Alliance on Race & Equity (GARE)
- [Racial Equity Toolkit: An Opportunity to Operationalize Equity](#), Government Alliance on Race & Equity (GARE)
- [Awake to Woke to Work](#), Equity in the Center, a Project of ProInspire

## Organize

**In your site context, how will the community and government learn, work, and be mutually accountable for using integrated data to inform, evaluate, and co-create structures, policies, practices, and narratives for equity?**

**Questions to guide your thinking should include:**

- How will community expertise be forefront throughout the data life cycle?  
What capacity will need to be developed to ensure that this occurs?
  - In planning
  - In data collection
  - In data access
  - In algorithms/use of statistical tools
  - In data analysis
  - In reporting and dissemination
- How will data use help communities interrogate systems, rather than just inform how to “treat” communities with additional services and programs?
- How will a racial equity lens be incorporated throughout the data life cycle?
- How will the culture, policies, practices, and expectations of the agency/collaborative shift to center racial equity?
- How will the agency/collaborative initially focus upon, continuously learn from, and sustain institutional-systems change?

## Operationalize

**What approaches will be most effective for integrated data infrastructure development and data use?**

- **For policy makers, agency, and community capacity building?**
- **For collaborating and power sharing among stakeholders?**
- **For centering community?**

### **1. Internal** considerations for agency/collaborative

- a. What work has been done with your policy leaders, agency directors, department staff, and front-line employees to prepare them for this work?
- b. Where will you start?
- c. Who is best prepared to lead this work, and what resources will they receive?

### **2. Internal AND external** considerations for agency/collaborative

- a. How will the agency/collaborative engage with community members, and in what manner?
- b. Who will hold the agency/collaborative accountable?
- c. How will the agency/collaborative ensure that the benefits of data integration outweigh the risks?

### 3. **Community** considerations

- a. How will community members/people whose data are in the system receive capacity-building? What structures will be implemented, including funding, training, and ongoing relationship building?
- b. What role does community organizing play in the design and use of data?
- c. What guides accountability and power sharing between community members/ participants and policy makers/agency?

## Our Approach

NORMALIZE	ORGANIZE	OPERATIONALIZE
<p><b>USE A RACIAL EQUITY FRAMEWORK</b></p> <p>Jurisdictions need to use a racial equity framework that clearly names the history of government in creating and maintaining racial inequities; envisions and operationalizes a new role; and utilizes clear and easily understood definitions of racial equity and inequity.</p> <p><b>COMMUNICATE &amp; ACT WITH URGENCY</b></p> <p>Despite the belief that change is hard and takes time, we have seen repeatedly that when racial equity is an urgently felt priority, change can be embraced and take place quickly. Building in institutional accountability mechanisms via a clear plan of action will allow accountability. Collectively, we must create greater urgency and public will to achieve racial equity.</p>	<p><b>BUILD ORGANIZATIONAL CAPACITY</b></p> <p>Jurisdictions need to be committed to the breadth and depth of institutional transformation. While the leadership of elected members and top officials is critical, changes take place on the ground, and infrastructure that creates racial equity experts and teams throughout local and regional government is necessary.</p> <p><b>PARTNER WITH OTHER INSTITUTIONS &amp; COMMUNITIES</b></p> <p>To achieve racial equity, local and regional government must work with a network of partners: institutions, business, education, philanthropy, among others, and center the work on impacted communities.</p>	<p><b>IMPLEMENT RACIAL EQUITY TOOLS</b></p> <p>Racial inequities are not random—they have been created and sustained over time. Inequities will not disappear on their own. Tools must be used to change the policies, programs, and practices that are perpetuating inequities, as well as used in the development of new policies and programs.</p> <p><b>BE DATA-DRIVEN</b></p> <p>Measurement must take place at two levels – first, to measure the success of specific programmatic and policy changes, and second, to develop baselines, set goals, and measure progress towards community goals.</p>

Source: [GARE Communications Guide](#), May 2018.

## Benefits, Limitations, & Risks

After assessing your agencies' readiness through the questions and concepts in the previous section, we encourage you to then consider specific examples of data sharing and integration. Because of the complexity inherent in civic data use, it is essential to carefully consider the benefits, limitations, and risks of each potential use.

### Benefits:

- Whole-person view:** When data are integrated across multiple sources, we get a more holistic view of the experiences and outcomes of children, households, and families, supporting asset- (rather than deficit- ) based approaches. Such views allow analysts to identify bright spots across communities, families, and individuals, and, ultimately, encourage investment in policies and programs that work.
- Scale:** Analysis using administrative data can include a whole population, rather than a sample, with longitudinal views and comparison groups readily available.
- Time & cost:** Reusing administrative data originally collected in the course of service delivery to answer important implementation and evaluation questions can be less time- and resource-intensive than collecting new data using surveys or other means.

### Limitations:

- Availability:** Since administrative data were originally collected by agencies for operational purposes, certain information will not be captured in these records.
- Quality:** Since administrative data are not collected for research purposes, data quality issues are common, including missing data, lack of data documentation, and questions related to reliability and validity.
- Access:** Many agencies do not have clear processes and procedures for sharing administrative data, which can make the process of gaining access difficult and time-consuming.

### Risks:

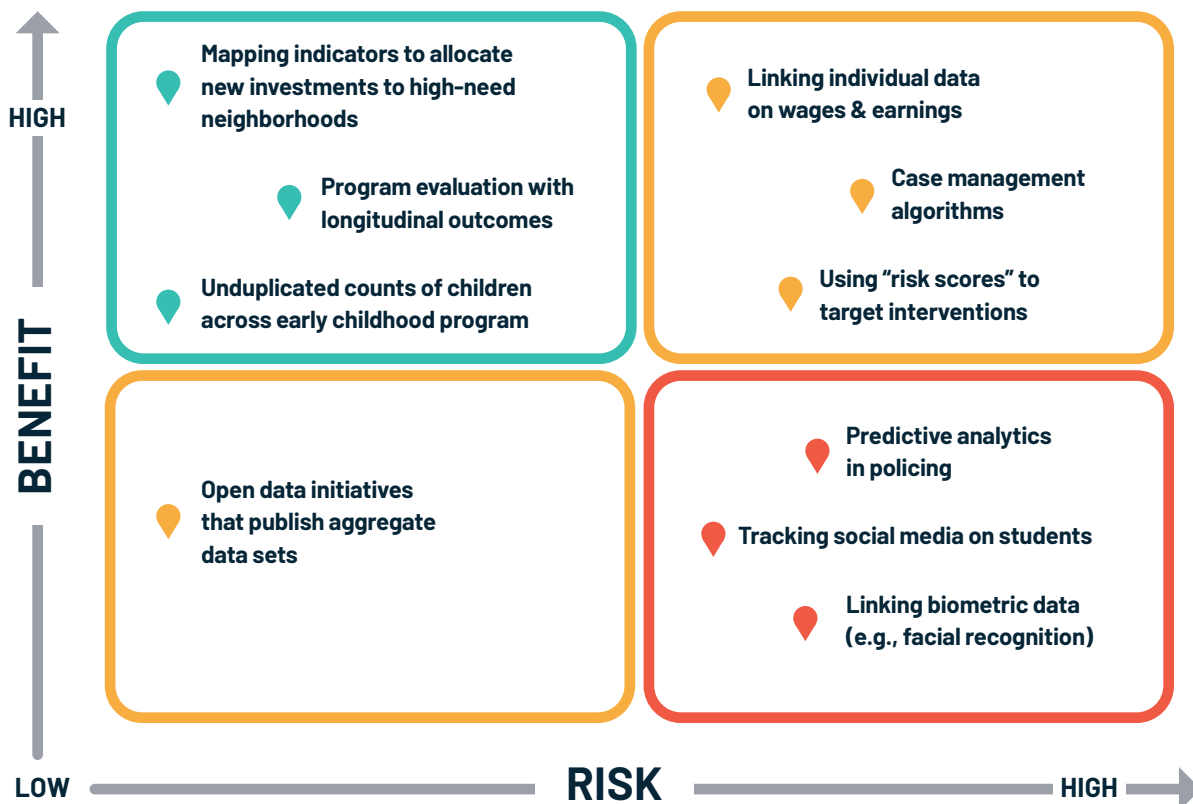
- Privacy disclosure:** The transfer of data includes the risk of data being accessed improperly, either by accident, through a security breach, or as a result of insufficient anonymization techniques. Such instances are rare with appropriate safeguards in place, but important to consider.
- Misuse of data for research and evaluation:** When data originally collected for administrative rather than analytic purposes are used for research, there is a risk that they may be misused or misinterpreted. Potential traps include the misuse of data as a result of an inappropriate analytic plan (e.g., using inferential statistics with an insufficient sample size), misuse of a variable (e.g., incorrectly assuming that "PRGENT" refers to date of program entry rather than to completed program entrance exam), or the inclusion of incorrect assumptions when explaining outcomes (e.g., explaining a reduction in out-of-school suspensions as being a positive indicator of improved climate, without knowing about a code of conduct revision that encouraged administrators to report suspensions differently).

- Replicating structural racism:** Since administrative data are collected during the administration of programs and services for individuals in need of social supports, the data include people who are disproportionately living in poverty, who, as a result of the historical legacy of race in America, are disproportionately Black, Indigenous, and people of color. Seeing data as race-neutral is inaccurate, and such views could lead to system-level data usage that unintentionally replicates structural racism.
- Harming individuals:** Certain individual-level uses of administrative data carry particularly high risks of causing personal harm. These include uses that provide case workers, service providers, teachers, law enforcement, etc., with personal information that could lead to biased treatments or punitive action, or lengthen system involvement.

While the particulars are important, we have identified broad categories of use on the risk vs. benefit matrix below.

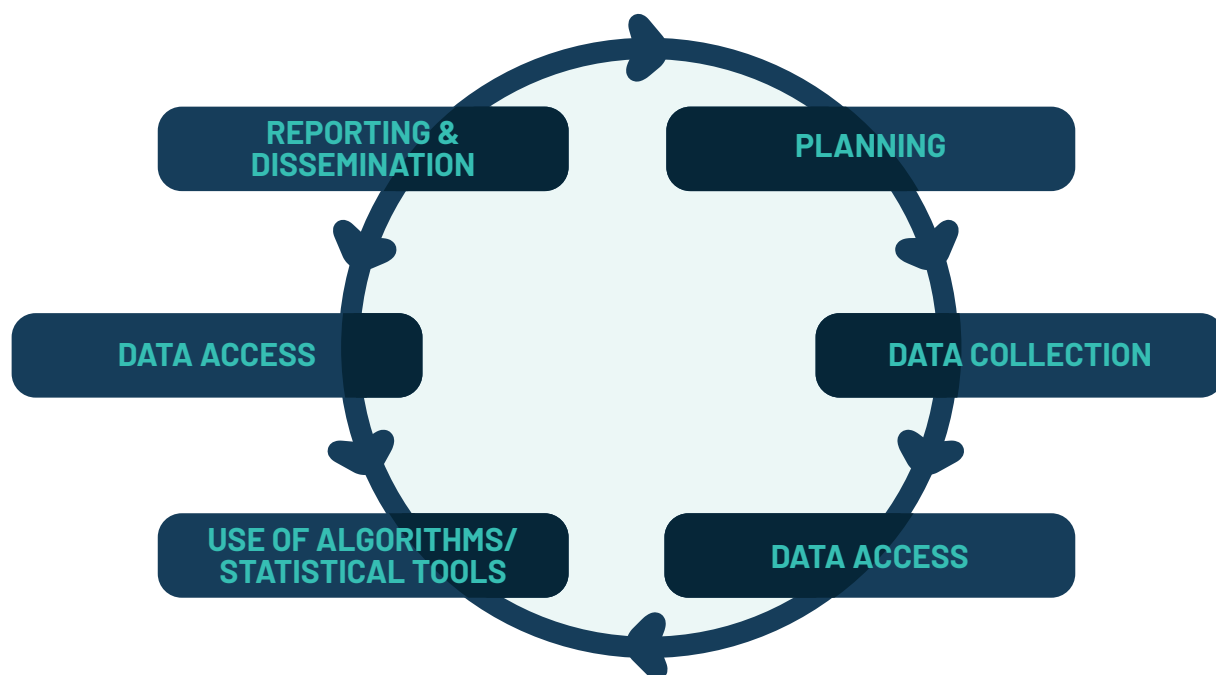
For example, projects that involve low risk and high benefit, such as a longitudinal program evaluation, indicator projects, or generating unduplicated counts across programs, are generally a good idea and an easy starting point for collaboration. Conversely, projects that are low benefit and high risk, such as [linking social media content with educational records](#) or [police surveillance using biometric data](#), should be considered with the utmost caution, and in some instances data sharing should not proceed.

**Examples included in this toolkit are not meant to be used as binary yes/no decision-making tools, but rather as guides for thinking about and talking through intended and unintended consequences of civic data use with a variety of voices and perspectives.**



## Centering Racial Equity Throughout the Data Life Cycle

Centering racial equity throughout data integration is not a single, discrete step, but rather an ongoing process at each stage of the data life cycle—planning, data collection, data access, use of algorithms and statistical tools, data analysis, and reporting and dissemination. Each stage presents new opportunities to bring a racial equity frame to data integration, as well as new challenges and considerations.



The following sections provide an overview of racial equity considerations throughout the data life cycle, examples of positive and problematic practices, and brief examples of Work in Action. The Work in Action examples highlight current site-based work taking place at each stage of the data life cycle. **Full-length versions of these examples of centering racial equity throughout the data life cycle are in Appendix I.**



## Racial Equity in Planning

Planning is the first stage of the data life cycle and includes all work to prepare for future stages, such as identifying stakeholders (see [Toolkit Activity 1](#)); convening work groups; articulating a mission or purpose for data integration; developing understanding of the local racial, social, and historical context (see [Toolkit Activity 2](#)); creating ethical guidelines for use; and developing a project plan. It is a common mistake to let the forward momentum and looming deadlines at the outset of a data project overshadow upfront equity work. However, bringing a racial equity frame to the planning stage will help set the standard for infusing racial equity throughout the rest of the data life cycle, making it a crucial first step.

### During this stage, first ask:

- Why is this work necessary?
- Who does the work benefit?
- How does it benefit the community at large?
- Who can the process/product harm?

These questions will help identify stakeholders, establish a project plan, and create a framework to center racial equity throughout the data life cycle.

Need more information to get started? Start with Chicago Beyond (2019), [Why am I always being researched? A guidebook for community organizations, researchers, and funders to help us get from insufficient understanding to more authentic truth.](#)

## Racial Equity in Planning: Positive & Problematic Practices

POSITIVE PRACTICE	PROBLEMATIC PRACTICE
Including diverse perspectives (such as community members with lived experiences and agency staff who understand the data) on planning committees	Using only token “representation” in agenda-setting, question creation, governance, or IRB review
Building capacity for researchers, administrators, and community participants to work together on agenda-setting	Using deadlines or grant deliverables as an excuse to rush or avoid authentic community engagement
Researching, understanding, and disseminating the history of local policies, systems, and structures involved, including past harms and future opportunities	Using only historical administrative data to describe the problem, without a clear plan of action to improve outcomes
Building data literacy among organizations and community members, which could range from light engagement through public activities like data “gallery walks” to more intense involvement, such as community-based participatory action research	Failing to manage expectations around what the data are capable of telling or how long it will take to see marked changes in data, actions, and outcomes
Establishing a common language and agreed upon sources and methods for reporting on community-based indicators	Failing to revisit indicator and outcome metrics regularly and revise when necessary
Clearly discerning who decides how to frame the problem or determine what questions to ask	Relying on academic institutions to frame the problem and research questions while failing to engage community-based organizations
Planning that includes the use of an asset; creating a framework that aims to clarify how to improve policy, services, and outcomes	Planning that includes the use of a deficit; creating a framework to describe outcomes
Lifting up the research needs of community to funders; helping shape funding strategy with funders to support community-driven research	Accepting grant/philanthropic funding for a project that is not a community priority or need

## Racial Equity in Planning: WORK IN ACTION

### Broward Data Collaborative by Sue Gallagher

Created in 2017, the Broward Data Collaborative (BDC) seeks to improve the outcomes of residents by integrating data from multiple sources to inform evaluation, research, and care coordination. The BDC consists of Broward County Public Schools, Broward County Human Services Department, Broward Behavioral Health Coalition, Florida Departments of Children and Families and Juvenile Justice, Early Learning Coalition of Broward, and the Children's Services Council of Broward County (which acts as the BDC backbone organization). In planning the BDC, the Children's Services Council of Broward County used a Community Participatory Action Research (CPAR) framework to operationalize their core values of equity and transparency and account for the historical and current structural racism perpetuated by the community's systems and organizations. The CPAR framework aims to help the BDC create racially equitable structures by involving system participants in the process of governance, research, evaluation, and solution creation to address the racial, economic, and social/spatial gaps that are all too common between the predominantly White researchers governing integrated data systems and the BIPOC using public services.

Through this planning process, the BDC has worked to provide a seat at the proverbial (and literal) table for youth and residents whose data are represented in the child-serving integrated data system. The BDC's recognition that their county-level systems produce racially disparate outcomes has allowed them to build nonhierarchical relationships between the community and professionals. In planning, the BDC is creating an integrated data system that allows them to share strengths-based stories about their community and community members and use data to co-create system and policy improvements.

### Tacoma Equity Index by Alison Beason

The City of Tacoma, Washington, created a Strategic Visioning Framework to guide the city's decision making through 2025. Collaborating with community members and groups during this planning process helped center themes of equity, opportunity, partnerships, and accountability in Tacoma's strategic vision. As a result, Tacoma established both the Equity and Empowerment Initiative and the Office of Equity and Human Rights to operationalize the city's strategic vision for service delivery around inclusivity and accessibility. Working with the Kirwan Institute of Race and Social Justice at The Ohio State University, the City of Tacoma created an Equity Index to measure social mobility in the city. Drawing from the concept of opportunity mapping, the Equity Index uses 20 indicators—aligned with the city's strategic goals—to measure opportunity and equity in the city. To date, the tool has been used to inform community conversations on equity and disparities and to guide policy makers' approaches to decision making and resource allocation.

## Racial Equity in Data Collection

Data collection is the process of gathering information to inform the study of a program, policy, or problem. Administrative data are typically collected for operational purposes, whereas other data collection efforts may arise from specific research or evaluation questions. Across agencies and programs, administrative data collection can look vastly different. For instance, data can be collected from participant intake paperwork, self-reported online applications, service payment records, participant surveys, or any other place where data have been collected to administer or evaluate programs or services.

While there are significant benefits to administrative data use, it is critical to understand that administrative data are collected for purposes other than research and therefore pose some potential risks. Most relevant for centering racial equity, administrative data are vulnerable to biases, inaccuracies, and incomplete or missing data, and most often include communities that are over-surveilled by government agencies. Implicit bias within data is commonplace and most often takes the form of selection bias (i.e., the individuals included in the data are not random or do not represent the intended population) or confirmation bias (i.e., data are used to confirm pre-existing beliefs). An equity lens considers these inherent vulnerabilities to the data collection process and how they can be reduced or contextualized appropriately in response. Remaining aware of one's own biases is an important first step.

**Intersectionality refers to the interconnected identities of an individual—including race, gender identity, nationality, sexual orientation, and disability. Collecting demographic data to better understand disparate impact based upon intersectionality can be an important step toward equity.**

"Intersectionality is simply a prism to see the interactive effects of various forms of discrimination and disempowerment. It looks at the way that racism, many times, interacts with patriarchy, heterosexism, classism, xenophobia—seeing that the overlapping vulnerabilities created by these systems actually create specific kinds of challenges. 'Intersectionality 102,' then, is to say that these distinct problems create challenges for movements that are only organized around these problems as separate and individual. So when racial justice doesn't have a critique of patriarchy and homophobia, the particular way that racism is experienced and exacerbated by heterosexism, classism, etc., falls outside of our political organizing. It means that significant numbers of people in our communities aren't being served by social justice frames because they don't address the particular ways that they're experiencing discrimination."

— Kimberlé Williams Crenshaw<sup>13</sup>

<sup>13</sup> Quoted in Guobadia, O. (2018), *Kimberle Crenshaw and Lady Phyll Talk Intersectionality, Solidarity, and Self-Care*.

## Racial Equity in Data Collection: Positive & Problematic Practices

POSITIVE PRACTICE	PROBLEMATIC PRACTICE
Adhering to data management best practices to secure data as they are collected—specifically, with carefully considered, role-based access	Assuming that programmatic staff (those most likely to collect data) have training in data management and data security
Including agency staff and community stakeholders in defining which data should be collected or reused	Inviting only researchers to identify data needs
Collaborating to develop a shared data collection agenda that is connected to practice, policy, and research	Collecting data that reinforces or confirms bias rather than informing practice and policy changes
Collaborating with agencies and community to generate a data development agenda—a plan for access and use of data that are needed to answer high-interest questions (e.g., expanding gender identity categories on a registration form; building support for digitizing eviction records)	Providing insufficient data labels (e.g., federal reporting in education has only seven race labels) or inconsistent categories across data sets (e.g., conflating race and ethnicity)
Working with staff to support equity-oriented data collection practices (e.g., programmatic staff to update a registration form, technical staff to update a “forced” field on a data entry platform)	Unwillingness to shift data collection practices based upon community feedback
Collecting only what is necessary to your context	Failing to consider which data carry an elevated risk of causing harm if redisclosed when determining which data to collect in your context (e.g., a housing program that collects resident HIV status)
Strong efforts to support metadata documentation, including key dimensions of metadata such as: <ul style="list-style-type: none"> <li>description</li> <li>technical specifications</li> <li>provenance</li> <li>rights</li> <li>preservation</li> <li>citation</li> </ul>	Failure to clearly identify, explain, and document data integrity issues, including data that are: <ul style="list-style-type: none"> <li>inaccurate</li> <li>undocumented</li> <li>unavailable</li> <li>incomplete</li> <li>inconsistent</li> </ul>
Including qualitative stories to contextualize quantitative data	Allowing quantitative data to “speak for itself” without context or discussion
Working with and developing flexible data systems that adapt to context, environment, or system changes	Working with and developing data systems that are static and offer limited access
Finding out why people “opt out” of providing data for surveys and other data collection efforts, and using their feedback to minimize harm in future data collection processes	Collecting data purely for surveilling groups marginalized by inequitable systems and BIPOC.

## Racial Equity in Data Collection: WORK IN ACTION

### Mecklenburg County Community Support Services

by Courtney Morton & Mary Ann Priester

In 2014, North Carolina's Mecklenburg County Community Support Services invested in two new positions tasked with improving data collection, data access, and community use of data: the Homelessness Management Information System (HMIS) Administrator and the Homelessness & Housing Research Coordinator. The HMIS Administrator works to expand community access to high-quality, accurate data on housing and homelessness by improving processes and procedures for collection, analysis, and implementation of best practices; the administrator supports Community Support Services' priority to end homelessness by ensuring that high-quality data are available to county and partner agencies. The Homelessness & Housing Research Coordinator is responsible for connecting community stakeholders with research and data to inform programming, policy, and funding decisions related to housing instability and homelessness. Collectively, the two new positions represent an investment by Community Support Services in Charlotte-Mecklenburg's commitment to collecting high-quality, useful, and timely data. The investment has led to important disaggregation across multiple intersectional categories, describing disparate impact of housing instability on Black residents—specifically Black women with children—therefore informing more equitable resource allocation.<sup>14</sup> The investment has also increased the capacity of community organizations to use data to inform funding, program, and policy decisions.

### Allegheny County, PA, Department of Human Services

by Shauna Lucadama & Jessica Ruffin

In 2009, after being approached by community members, Pennsylvania's Allegheny County Department of Human Services (DHS) began a department-wide effort to look at ways to provide culturally responsive, affirming services to the LGBTQ communities involved with DHS. At the time, gender and sexuality were not openly discussed in regard to DHS practices. In 2013, through a partnership with The Center for the Study of Social Policy, DHS began the work of piloting guidelines for collecting data related to sexual orientation, gender identity, and gender expression (SOGIE) for children and youth involved with the child welfare system. The steering committee created to lead this effort ultimately recommended updating the department's case management system to include SOGIE data, implement standards of practice related to SOGIE, and provide training and support to case workers implementing these new practices.

Before collecting new data, DHS had to address privacy and data security concerns surrounding youth SOGIE data, the implications of data being shared with external stakeholders, and the complexities and costs of updating an IT system. Additionally, DHS engaged with IT staff to ensure they knew the importance of these inclusive changes in order to mitigate any harm during the design process. A field test of the guidelines showed the broader child welfare community the importance of SOGIE data in improving services to the LGBTQ communities. Now, these guidelines are standard practices in providing services to families.

<sup>14</sup> See <https://mecklenburghousingdata.org/>

# Racial Equity in Data Access

Data access generally refers to practices regarding who can securely obtain, view, or use data; when it can be used; and for what purpose. We recommend that agencies and community-based organizations use a multi-tiered approach to data access that first begins with a clear delineation of practical and legal data availability.

OPEN DATA	RESTRICTED DATA	UNAVAILABLE DATA
Data that can be shared openly, either at the aggregate or individual level, based on state and federal law. These data often exist in open data portals.	Data that can be shared, but only under specific circumstances with appropriate safeguards in place.	Data that cannot or should not be shared, either because of state or federal law, lack of digital format (paper copies only), or data quality or other concerns.

The determination of which data are open, restricted, or unavailable can have significant implications for equity. For example, the intended or unintended consequences of a data release may disproportionately impact some individuals or communities more than others. Alternatively, the impact of not releasing data that are needed to understand or address a community-based problem may also be disproportionately felt.

**Open datasets** are most often available online, through search queries, static PDFs, CSV files, and front-end data visualization tools (such as dashboards). Whether open data supports or undermines racial equity is dependent upon the source, purpose, presentation, and context of the open data.

## Do the following examples of open data center racial equity?

- Salaries of public employees, [searchable by name](#)
- "Teacher Effectiveness" scores, [by teacher and school](#)
- [Housing market indicators](#)
- [County-based health disparities](#)
- [Traffic stops by race](#)
- [211 calls](#)
- [Voting records](#)
- [Daily "inmates" in custody<sup>15</sup>](#)
- [Arrest records](#)

<sup>15</sup> Across all populations, person-centered language is best practice, including using the term "person who is incarcerated" rather than "inmate." In this context we are linking to a data source titled in this way. See [Underground Scholars Initiative](#).

While each of these open data examples is permissible by law, there are a range of perspectives on whether and how these data should be made available for public use. To center racial equity, it is important that any release of open data be carefully considered from a range of perspectives—including agencies, individuals, and communities—and that data be interrogated within the racial, historical, and sociocultural context. For example, in the list above, traffic stops by race could be a data-informed tool used to push for change in policing, while these same data could also be used as justification for increased surveillance, as exposure to extreme disparities can cause people to become more, not less, supportive of the very policies that create those disparities.<sup>16</sup>

Whereas open source data can be freely accessed, used, and shared by anyone, restricted data are protected from public access by federal and state privacy laws, regulations, and statutes.

**Restricted data:** If you are a caseworker, then having access to individual records is an important aspect of your role. Alternatively, if you are a stakeholder external to the agency, having access to such restricted data in an identifiable format is typically not needed. If a need exists, formidable challenges arise as a result of the logistical mechanisms needed for secure access. In this instance, access is dependent upon a range of factors:

- **Why** you want to access the information (e.g., research, evaluation, or another purpose)
- **What** type of information you want to access (e.g., aggregate information or identified records)
- **Who** you are (e.g., agency analyst, independent evaluator, or concerned citizen)
- **How** you will share the information (e.g., anonymized findings via dashboard, aggregate reporting in a static document, publicly released names)

**Unavailable data** cannot or should not be shared for a variety of reasons. For some, such as HIV status or drug treatment, access by external users is prohibited by law. Other sources may be unavailable because they are not yet digitized (e.g., eviction records) or have substantial data quality issues. There are also data that are deemed too risky for any form of release in case of redisclosure, such as information on domestic violence.

Categorizing data as open, restricted, and unavailable is a large task but one that is important for centering racial equity. Both restricting and opening access to data can lead to equity concerns, and risks and benefits should be carefully considered.

<sup>16</sup> Hetey, R., & Eberhardt, J. (2018)



## Racial Equity in Data Access: Positive & Problematic Practices

POSITIVE PRACTICE	PROBLEMATIC PRACTICE
OPEN DATA	
Open data that have been identified as valuable through engagement with individuals represented within the data	Ongoing open data that is based upon problematic indexes or algorithms, with a history of discriminatory impact on communities (e.g., release of “teacher effectiveness scores” and “school report cards”)
Clear data release schedules and information on where to go and how to access data once they are released	Releasing data that can be re-identified (e.g., data released by small geographies may be identifiable by local residents)
Clear processes for submitting a request to agency for making data open, including how requests will be evaluated	
RESTRICTED DATA	
Adhering to data management best practices for data access, e.g., secure transfer of data (encryption in transit and at rest); clear data destruction parameters, if applicable, following use	Assuming that data management best practices are being followed without explicit protocols and oversight in place
Utmost care given to de-identification and anonymization of data prior to release	Releasing data that can be re-identified (e.g., data that have not been properly anonymized or include aggregate or subgroup data without suppressing small cell sizes)
Accessible data request process with clear policies and procedures for submitting a request and how requests are evaluated	Unwillingness to release data, or limiting access to researchers or individuals with an “in”
Creating, using, and sharing high-quality metadata to inform requestors about what data are available	Putting the onus solely on requesters to know which data can be requested
Free or low cost for data request, with costs clearly communicated (e.g., based on hourly rate or a flat fee)	Significant or inconsistent costs associated with requests for data access
Supporting directory-level data sharing among agencies for referral coordination and to lessen administrative burden on clients	
UNAVAILABLE DATA	
Clear documentation of why data are unavailable (e.g., specific statute, legislation, data quality explanation, data are not digitized, undue burden in data preparation)	Refusal to release data when release is permissible and would not pose an undue burden
	Organizations are unable to access their own data as a result of third-party contractors owning or controlling the data

## Racial Equity in Data Access: WORK IN ACTION

**Examples of Work in Action for data access fit into one or more of these categories:**

1. Transparency of process
2. Track record of providing access to a diverse range of stakeholders
3. Preservation of privacy

### KYSTATS by AISP with contributions from Jessica Cunningham

The Kentucky Center for Statistics (KYSTATS) is a state government-backed organization that collects and links education and workforce data for use by policy makers, practitioners, and the public. The KYSTATS website provides a public outlet for transparency on data collection, access, and use. The website hosts easily accessible information on all data that are collected, stored, and maintained. The [KYSTATS "Data Access and Use Policy"](#) provides context for how agencies, residents, and providers access data based on their user type. Published flow charts clearly articulate the data request process and response timeline. KYSTATS uses a cost-recovery model allowing service providers affordable access to the data that they need. Additionally, the [KYSTATS data use dictionary](#) aims to close the gap between data professionals and the public by providing context around what data are and are not available for use. KYSTATS also demonstrates a deep commitment to data privacy, with clear anonymization protocols that align with industry best practices to avoid redisclosure. This dual commitment to both privacy and access is important for supporting equitable data use and mitigating harm for individuals represented within the data.

### Allegheny County Department of Human Services Client Experience Analytics Unit by AISP with contributions from Samantha Looney, Brian Bell, Julia Reuben, Shannon Flynn, & Jamaal Davis

Allegheny County's Department of Human Services (DHS) in Pennsylvania is improving community access to data through the creation of a client portal. The portal hosts school district, housing, income support, child welfare, and other individual-level data housed in the county's data warehouse. The portal will allow DHS clients to access their personal data, coordinate their own care, and provide feedback on the quality of support they receive from public programs. Additionally, network providers will be able to access integrated client data, enabling them to provide more holistic care. The new system is being developed from community conversations, interviews, and a prototyping process with DHS program participants. Allegheny County DHS believes in community members' right to their own information. The client portal will allow residents to review their individual records generated through the receipt of county and community-based services. The community-centered design process supports equity in data access to ensure that the portal reflects community needs and benefits both clients and service providers.

### Birth through Eight Strategy for Tulsa (BEST) / Asemio Privacy-preserving Record Linkage Pilot by Jessica England & Dan Sterba

In 2019, eight organizations entered into a formal data collaborative in Tulsa, Oklahoma, aiming to identify and address community-level challenges regarding race, equity, and service overlap. The collaborative, known as the Birth through Eight Strategy for Tulsa (BEST), includes 12 years of data from county government, local government, nonprofit, private sector, and philanthropic organizations that represent 32 programs and 220,000 unique individuals. BEST piloted a new platform developed

by Asemio<sup>17</sup> utilizing privacy-preserving record linkage that supports data integration while keeping individual and organizational data private and secure. The platform's use of cryptographic and record linkage technology for community analytics allows researchers to integrate data more quickly, at lower cost, while enhancing privacy for individuals and organizations. The process, driven by agency leaders incentivized by access to more robust information and improved service delivery, produced results in less than two months while improving ethical and security protections.

## Racial Equity in Algorithms/Statistical Tools

Automated decision-making algorithms are commonplace in private industry and are becoming increasingly common across the public sector as well. Algorithms, or “series of instructions written and maintained by programmers that adjust on the basis of human behavior,”<sup>18</sup> are housed in statistical tool packages (e.g., SQL, R, SAS) and incorporate statistical tests and logic in order to automate decision making.

There is no such thing as a race-neutral algorithm, since algorithms reflect the biases of those who create them and the data used in their processes. That said, there are strategies and tools that can and should be used to ensure transparency, assess algorithmic bias, and determine the potential positive and negative consequences of applying an algorithm in practice.

The five guidelines (summarized below) from [Fairness, Accountability, and Transparency in Machine Learning](#)<sup>19</sup> can be used as a starting place to inform the development and use of algorithmic tools in ways that are accountable to the public.

- 1. Responsibility:** Create clear channels for communication about potential adverse impacts of algorithms and name specific individuals tasked with addressing these impacts
- 2. Explainability:** Ensure that algorithmic decisions and the data driving decisions can be explained to end-users and stakeholders in nontechnical terms
- 3. Accuracy:** Identify, log, and explain sources of error and uncertainty so that intended and unintended consequences can be anticipated and planned for
- 4. Auditability:** Enable third parties to monitor and evaluate algorithmic decisions
- 5. Fairness:** Ensure that algorithmic decisions do not create discriminatory or unjust impacts

In order to formalize a commitment to the public and to these principles, it is recommended that those developing and using algorithms draft a [Social Impact Statement](#) that describes how these five principles (or others relevant to the local context) will be operationalized in practice. For example, efforts committed to centering racial equity should, at a minimum, involve a variety of stakeholder groups in the continuous evaluation of the differential impacts, risks, and benefits of the algorithm's use. A number of tools and frameworks may be helpful in this process, including [Algorithmic Fairness: A Code-based Primer for Public-sector Data Scientists](#),<sup>20</sup> the [Ethics & Algorithms Toolkit](#),<sup>21</sup> the [Data Nutrition Project](#) prototype and white paper,<sup>22</sup> and [Model Cards](#).<sup>23</sup>

<sup>17</sup> See [asemio.com](https://www.asemio.com)

<sup>18</sup> Benjamin, R. (2019)

<sup>19</sup> Diakopoulos, N., et al. (2017)

<sup>20</sup> Steif, K. & Goldstein, S. (2019)

<sup>21</sup> Anderson, D., et al. (2018)

<sup>22</sup> Holland, S., et al. (2018)

<sup>23</sup> Mitchell, M., et al. (2019)

## Racial Equity in Algorithms and Statistical Tools: Positive & Problematic Practices

POSITIVE PRACTICE	PROBLEMATIC PRACTICE
Involving diverse stakeholders, including specific community advisory boards, in early conversations about the purpose of an algorithm prior to development and implementation	Developing and implementing algorithms for human services without stakeholder involvement or alignment across multiple agencies
Determining responsibility for oversight of algorithm development and implementation, with clear communication channels for input	Inadequate opportunities for community feedback regarding algorithm development and implementation
Mandatory impact assessments that involve thoroughly thinking through potential intended and unintended consequences	Failure to think through intended and unintended outcomes
Clearly identifying and communicating potential benefits and risks to stakeholders	Implementing an algorithm with no clear benefit to individuals included in the data
Human-led algorithm use (i.e., human can override algorithm at any point in process)	Elevating algorithmic decision making over judgment of seasoned practitioners; no human involvement
Transparency regarding what data drive the algorithm and how, e.g., description of design and testing process, list of factors that the tool uses, thresholds used, outcome data used to develop and validate the tool, definitions of what an instrument forecasts and for what time period	Use of a “black box” or proprietary algorithm that does not allow for transparency or replication
Efforts to improve the quality of data included within the algorithm, including efforts to balance the use of risk and protective factors	Use of data with data integrity issues or “dirty” data that reflect bias in data collection (resulting in garbage in/garbage out)
Using “early warning” indicators to provide meaningful services and supports to clients	Using “early warning” indicators for increased surveillance, punitive action, monitoring, or “threat” amplification via a risk score
Using multiple measures of validity and fairness. e.g., testing of metrics that center racial equity such as false positives/negatives across race and gender	Use of biometric data (specifically facial recognition)

## Racial Equity in Algorithms and Statistical Tools: WORK IN ACTION

### Allegheny County Child Welfare Office by AISP with contributions from Katy Collins

Allegheny County, Pennsylvania's Family Screening Tool (AFST) is a predictive analytic algorithm for child welfare that has received a great deal of study<sup>24</sup> and national press, including influential critiques.<sup>25</sup> While we remain concerned about the use of algorithms based upon underlying data that are rich with systemic bias and a history of discriminatory impact, we acknowledge Allegheny County's notable efforts in following best practices of algorithm development, including the five principles summarized above: responsibility, explainability, accuracy, auditability, and fairness.

The algorithm is powered by the county's [long-established integrated data system](#),<sup>26</sup> which provides child welfare caseworkers with a risk assessment score during the maltreatment hotline screening process. The risk assessment score, calculated based on data points from eight county databases, provides caseworkers with a tool to engage in data-informed decision making, while prioritizing human judgment during the process. In building the algorithm, Allegheny County provided space for legal experts, academics, community activists, and former foster youth to provide feedback and input into the system's construction and use. Additionally, the county commissioned an independent ethical review and process and impact evaluations before implementing the AFST. Since the algorithm's inception, the county has seen a reduction in observed racial disparity among cases screened for further investigation. Importantly, families with low-risk cases are less frequently subjected to intrusive investigations, while high-risk cases are being investigated at higher rates.

### Automating.NYC by AISP with contributions from Deepra Yusuf, Elyse Voegeli, Akina Younge, & Jon Truong

New York City convened the Automated Decision Systems Task Force to assess the use and proliferation of automated decision systems (ADS) across social and city services in May 2018. The task force, which was mandated under Local Law 49, required participants develop a series of recommendations on the use of ADS and would include community outreach and input as part of this process. In January 2019, four graduate students—Akina Younge, Deepra Yusuf, Elyse Voegeli, and Jon Truong—constructed the website [Automating.NYC](#) to help make conversations about ADS more accessible to community members before the first public engagement forum hosted by the Task Force in late April. It aimed to support community members in asking informed questions about how automated decisions can and do contribute to unjust systems, and with the hope that future systems are built to benefit them. Members of the group were all previously involved in efforts to promote algorithmic justice and had been accumulating knowledge on NYC-specific ADSs since fall of 2018.

To build the website, they used information from investigative reporting, research, and open data, and worked with New York agencies to develop case studies. Accessible language and formatting, as well as a series of examples across different social services systems, allows for a broad audience with varying experience or awareness of ADSs. Nontechnical topics are used to describe technical terms and concepts, such as buying avocados to explain a decision tree algorithmic (adapted from Cathy O'Neil's book *Weapons of Math Destruction*), with engaging and interactive activities. Individual stories always accompany technical explanations, thereby merging theoretical explanation of automated decision making with its application and the felt impact on the local population.

<sup>24</sup> Kingsley, C. (2017)

<sup>25</sup> Eubanks, V. (2018)

<sup>26</sup> See AISP, [Network Sites](#), Allegheny County, PA

Most notably, the “toy algorithm” built using data from the NYC Open Data Portal offers users the opportunity “to help an imaginary fire department predict which buildings are at high risk for fire” using a series of variables; it demonstrates how algorithms, though driven by data points, are always influenced by the human decision making about what goes into them. Finally, the site concludes with action steps for users to take and apply after they close the browser window.

## Racial Equity in Data Analysis

Data analysis is the stage at which available data are explored in order to develop findings, interpretations, and conclusions. Data analysis can be as simple as calculating descriptive statistics, such as counts of program participants or the percentage of participants who achieved a certain outcome. Analysis can also include creating complex models that predict participant behavior or outcomes based on certain inputs, and these considerations are broadly discussed in the previous section.

Incorporating a racial equity lens during data analysis includes incorporating individual, community, political, and historical contexts of race to inform analysis, conclusions, and recommendations. Solely relying on statistical outputs will not necessarily lead to insights without careful consideration during the analytic process, such as ensuring data quality is sufficient and determining appropriate statistical power. Disaggregation of data is also a series of tradeoffs. Without disaggregating data by subgroup, analysis can unintentionally gloss over inequity and lead to invisible experiences. On the other hand, when analysts create a subgroup, they may be shifting the focus of analysis to a specific population that is likely already over-surveilled.

Given the complex series of decisions inherently involved in the process of centering equity within data analysis, collaborative and iterative work with strong participation from a variety of stakeholders is critical.

## Racial Equity in Data Analysis: Positive & Problematic Practices

POSITIVE PRACTICE	PROBLEMATIC PRACTICE
Using participatory research to bring multiple perspectives to the interpretation of the data	Describing outcomes without examining larger systems, policies, and social conditions that contribute to disparities in outcomes (e.g., poverty, housing segregation, access to education)
Engaging domain experts (e.g., agency staff, caseworkers) and methods experts (e.g., data scientists, statisticians) to ensure that the data model used is appropriate to examine the research questions in local context	Applying a “one size fits all” approach to analysis (i.e., what works in one place may not be appropriate elsewhere)
Correlating place to outcomes (e.g., overlaying redlining data to outcomes)	Leaving out the role of historical policies in the interpretation of findings
Using appropriate comparison groups to contextualize findings	Making default comparisons to White outcomes (e.g., assuming White outcomes are normative)
Employing mixed methods approaches when developing the analytic plan, including purposefully seeking out qualitative data (interviews, focus groups, narrative, long-form surveys) in conjunction with quantitative administrative data to better understand the lived experience of clients	Using one-dimensional data to propel an agenda (e.g., use of student test scores in isolation from contextual factors such as teacher turnover, school-level demographics)
Disaggregating data and analyzing intersectional experiences (e.g., looking at race by gender)	Disregarding the individual or community context in the method of analysis and interpretation of results
Empowering professionals and community members to use data to improve their work and their communities	Analyzing data with no intent to drive action or change that benefits those being served



## Racial Equity in Data Analysis: WORK IN ACTION

### #ChangeFocusNYC by Sarah Zeller-Berkman

In 2016, the Administration of Children's Services (ACS) in New York City partnered with the Department of Education to examine a series of indicators common to both departments' datasets. The project brought together agency leads, community activists, and academics to form a community engagement committee tasked with ensuring that nonagency voices helped shape the analysis. The committee held meetings with individuals working in schools, community-based organizations, and health clinics to identify indicators related to education and child abuse, prevention, and detention that would be analyzed by researchers. Fifteen youth were chosen to partner with academics at the Youth Studies Program at the CUNY School of Professional Studies to design and implement a participatory action research project formally titled #ChangeFocusNYC.

#### #ChangeFocusNYC set out to answer two main questions:

- What are the experiences of NYC youth ages 14–21 who have dealt with multiple city agencies in their lives?
- What are policy and/or programmatic recommendations that could benefit youth ages 14–21 who are dealing with multiple agencies in their lives?

Youth investigators were involved in all phases of the research project and were essential contributors during development of the analytic plan. Collaboratively generated answers to the research questions will help ACS work toward creating a system in which young people are continuously engaged in shaping the institutions that impact their lives.

## Racial Equity in Reporting and Dissemination

Reporting and dissemination is the final stage in the data life cycle, where findings from the data are communicated internally or externally. This involves strategic consideration of the audience and the mode of dissemination that most effectively conveys the information. While a static PDF report (like this one) currently remains the most common way to share findings, there are countless ways to communicate information, as demonstrated by the Work in Action sections featured throughout this toolkit. These include briefs, interactive documents, websites, dashboards, social media content, data walks, posters, briefs, and infographics.

Regardless of the form, content geared toward the public should avoid jargon that may be otherwise appropriate for internal program staff or academic audiences, while also using person-centered language<sup>27</sup> and translating materials into languages most applicable to your community context. Unsure about your accessibility? Here is a [checklist for plain language](#)<sup>28</sup> and a [plain language thesaurus](#).<sup>29</sup> You can also check the [readability of your text](#).<sup>30</sup> Finer details, such as color palettes suitable for color blindness or printing in grayscale, are also important when communicating to larger, external audiences.

<sup>27</sup> Hyams, K., et al. (2018)

<sup>28</sup> Plain Language Action and Information Network (n.d.)

<sup>29</sup> National Center for Health Marketing (2007)

<sup>30</sup> Readability Formulas (n.d.)



Reporting on data work can include formal, written documents, infographics and data visualizations, website materials, press releases or news articles, and even speeches. Across all of these mediums, centering racial equity means paying attention to which data are highlighted and how they are framed, as well as the readability and accessibility of the communication method.

## Racial Equity in Reporting and Dissemination: Positive & Problematic Practices

POSITIVE PRACTICE	PROBLEMATIC PRACTICE
Creating a range of products to communicate findings across a wide variety of audiences via both online and offline methods of dissemination	Creating one output that is inaccessible to general audiences (e.g., a 100-page static report, content behind a paywall)
Developing differentiated messaging for different audiences that considers the appropriate level of detail and technical jargon, language, length, format, etc.	Using intentionally dense language with low readability, especially for non-native language learners
Reporting data in an actionable form to improve the lives of those represented in the data (e.g., analyzing food purchase data to identify food deserts and guide development of grocery stores)	Reporting data that are not actionable or that are intended to be punitive (e.g., analyzing food purchase data to remove recipients from Temporary Assistance for Needy Families [TANF] or other benefits)
Providing public access to aggregate data (e.g., dashboards, routine reports)	Putting materials solely online, particularly behind a paywall
Acknowledging structural racism or other harms to communities that are embedded in the data	Attempting to describe individual experiences with aggregate or “whole population” data without analyzing disparate impact based on race, gender, and other intersections of identity
Including stories as a complement to quantitative findings in order to better contextualize the lived experience represented by the numbers	Allowing the data to “speak for itself” without context or discussion
Providing clear documentation of the data analysis process along with analytic files, so that others can reproduce the results	Obscuring the analytic approach used in a way that limits reproducibility
Conducting impact analyses multiple times during the project (e.g., at the beginning, middle, and end). Asking the core question: does this work mitigate, worsen, or ignore existing disparities?	Disregarding how findings will impact individuals or communities

## Racial Equity in Reporting and Dissemination: WORK IN ACTION

### DataWorks NC by Libby McClure & John Killeen

DataWorks NC is a nonprofit, community-data intermediary in Durham, North Carolina. One of DataWorks NC's main projects is the Neighborhood Compass, a public-facing mapping tool that reports administrative data and other publicly available information. Motivated by community conversations in Durham around the state of evictions, DataWorks NC created a series of events titled, Who Owns Durham? These events disseminated information through "gallery walks," participatory exhibits, and community conversations, enabling tenants and community groups to generate ideas surrounding displacement prevention.

These bidirectional reporting practices have allowed for collaboration with community members and resulted in the creation of a strategy document, "Quick Takes," and webpage for use by organizers, tenants, and other community groups addressing housing inequities.

### Metriarch Data Lookbook and Lady Charts Event by Emma Swepston, Laura Bellis, & Brandy Hammons

In 2018, the Take Control Initiative in Tulsa, Oklahoma, created a statewide women's health data collaborative. The data collaborative, known as Metriarch, aims to normalize and broaden conversations around women's health in Oklahoma through data storytelling, resource curation, and interactive outreach events. Metriarch has taken an innovative approach to analyzing and disseminating data in a way that educates and engages community members, service providers, and policy makers.

The Metriarch Lookbook's narrative approach to data storytelling makes their analysis understandable and usable for a wide range of statewide collaborators. Metriarch's Lady Charts data jams provide an opportunity for the community to engage in conversation around data to better understand what they can do to improve community outcomes. The data jams bring together diverse perspectives for meaningful dialogue on what data represents and what/who is still missing from the narrative. Relatedly, Metriarch has created a legislative tool that tracks pending legislation in the Oklahoma state legislature. The tool provides a platform to transform data into advocacy and activism, creating more equitable outcomes for women in Oklahoma.

### City of Asheville, Mapping Racial Equity by Christen McNamara & Kimberlee Archie

Asheville, North Carolina, has seen major increases in population, tourism, and economic activity over the past decade. These increases have had unintended negative consequences for many low-income residents and residents of color, leading to widespread gentrification and displacement. The City of Asheville Office of Equity & Inclusion recognizes the local government's role in creating and maintaining disparities experienced by people of color and people living in poverty, and is working toward shifting their operations to redistribute power and improve community conditions. Recently, they partnered with the city's GIS office to create a story map, "[Mapping Equity in Asheville](#)." Linking racial demographics over time to location and publishing the results online in a user-friendly format has allowed Asheville residents to better understand the connection between racialized policies and physical location, particularly in regards to redlining practices. The story map provides valuable information to all levels of government and community members to help inform policy and program development and resource allocation.

## What Next?

While data sharing and integration are essential tools for data-informed decision making within communities, to date, racial equity has been a peripheral consideration, at best. **We envision the use of data integration as a means to confront racism, expose injustice, act on our shared values, and elevate lived experience to co-create and implement new and innovative equity structures and narratives.** We aim to use collective knowledge and skills to advance government transparency and accountability in data use, which is critical to building trust, community well-being, and improved outcomes.

We encourage you to use the resources, stories, and activities shared here to shift your awareness and practice toward actualizing a data infrastructure that centers racial equity. We believe our work together can support sustained investment in government and community capacity for collaboration around ethical data use.

We encourage you to share your experiences in doing the important work of centering racial equity throughout the data life cycle. Contact us at [www.aisp.upenn.edu](http://www.aisp.upenn.edu), so we can learn alongside stakeholders and continue to share best practices and Work in Action.

# **Toolkit Activity 1:** **Who Should Be at the Table?**

*This activity offers guidance on identifying, assembling, and gathering around a table of diverse stakeholders for meaningful engagement, and should be the beginning point for data integration work that centers racial equity.*

Meaningful engagement occurs when stakeholders are able to have true influence over the design and direction of a data integration effort and its use cases. Tokenism and performative inclusion will undercut trust and relationships, particularly among groups historically underrepresented among and marginalized by dominant decision-making processes.

Members of the local community and those most represented in the data must always have seats at the decision-making table. Diverse stakeholder engagement is essential to informed, ethical planning and a key component to establishing trust and strong relationships throughout the data life cycle.

## Identifying the table

Resource availability will help inform the size, location, and design of your table. Start by assessing your financial and organizational capacity for engagement in order to set realistic goals and expectations for your effort.

The following questions and considerations have been adapted from AISP and Future of Privacy Forum's [Nothing to Hide: Tools for Talking \(and Listening\) About Data Privacy for Integrated Data Systems](#).<sup>31</sup>

### Timeline:

- How frequently will stakeholders be engaged?
- How much education or level-setting will you need to do with stakeholders before each engagement activity?
- How much notice or time to mobilize (if any) do potential participants need before each engagement activity?

### Budget:

- How many participants do you anticipate inviting?
- Will engagement activities need funding?
- What materials or resources need to be gathered or created/published/printed?
- Will participants need to be engaged or reimbursed for travel or other expenses?
- Will any participants receive compensation?

<sup>31</sup> Future of Privacy Forum & AISP (2018)

**Capacity for change:**

- What is the current decision-making environment for your organization? Who has final authority?
- How committed is internal leadership to meaningful/active stakeholder engagement (i.e., will input have material impact on outcomes)?
- What aspects are open to stakeholder influence?
- What aspects are non-negotiable?

**Staff capacity:**

- Do you or partners have staff . . .
  - With meeting facilitation experience?
  - With participatory engagement experience?
  - With subject matter expertise in racial equity?
  - With communications and public outreach experience?
  - Who can offer technical and logistical support for engagement activities, including events?
  - Who are from the local community?
  - Who hold marginalized identities?

Stakeholder engagement must be inclusive of diversity in race, language, culture, socioeconomic status, and ability. **Keep the following in mind as you plan your engagement activities:**

- Language and literacy, such as using bilingual facilitators or providing translators.
- Food, such as providing halal, kosher, or vegan meals.
- Location and transportation, such as meeting near public transit or spaces familiar to traditionally marginalized or underserved populations.
- Time, such as hosting multiple engagements or hosting outside of traditional work hours, or accommodating prayer times for religious participants.
- Childcare, such as providing childcare during in-person meetings.
- Incentives, such as providing free food and drinks at meetings, or compensating participants for their time and knowledge.
- Appeal, such as increasing use case relevance and impact on particular community groups.
- Accessibility (physical and digital), such as providing ADA-compliant physical and digital spaces or providing assistive technologies for those who need it.

## Assembling the table

Stakeholder mapping is an opportunity to evaluate what groups, perspectives, and knowledge are or should be represented at the table. The following three-step process outlines how to identify, assess, and prioritize the people with whom you will engage.

Remember, knowledge and expertise do not refer to specific degrees, titles, or accolades. Rather, an equitable approach to engagement acknowledges that expertise comes from a variety of sources—such as education, work experience, lived experiences, community participation, and service participation. This will inform how you think about stakeholder engagement.

### 1. Identify your stakeholders.

- Core stakeholders, whose engagement is central to data infrastructure
  - Data owners and contributors (directly contributing, or facilitating access)
  - Funding sources (government, private foundations, other)
  - Public agency leadership and key elected officials
- Other direct stakeholders, whose engagement can help facilitate (or impede) data sharing success but who are not in the core group
  - Data users (researchers, advocacy groups)
  - Technical experts (legal, data technology, security, research methods, fiscal)
  - Members of communities marginalized by inequitable systems
  - Advocates for vulnerable populations and communities
- Other stakeholders, who can broaden interest of data sharing and deepen its constituencies
  - Business groups
  - Good government groups
  - Other citizen and public interest groups

#### Consider these questions throughout the process to ensure no one is overlooked:

- Who will be representing the interests of the individual community members whose administrative data are being used?
- Which people or organizations will be affected by the results of your data sharing effort now and in the future?
- Which people or organizations are influential on this issue at the local, state, national, or international level?
- Who is influential within your particular area, community, or organization?
- Who can obstruct a decision if they are not involved (individuals, funders, political leaders, oversight groups, etc.)?
- Who has been involved in this issue in the past?
- Who has not been involved in past engagements, but should have been?
- Are there any barriers to engagement that may be/have been deterring some stakeholders?

## 2. Assess your stakeholders' interests.

Here, it is important to think through why (or why not) and in what ways each group of stakeholders is likely to advance or impede efforts around data sharing.

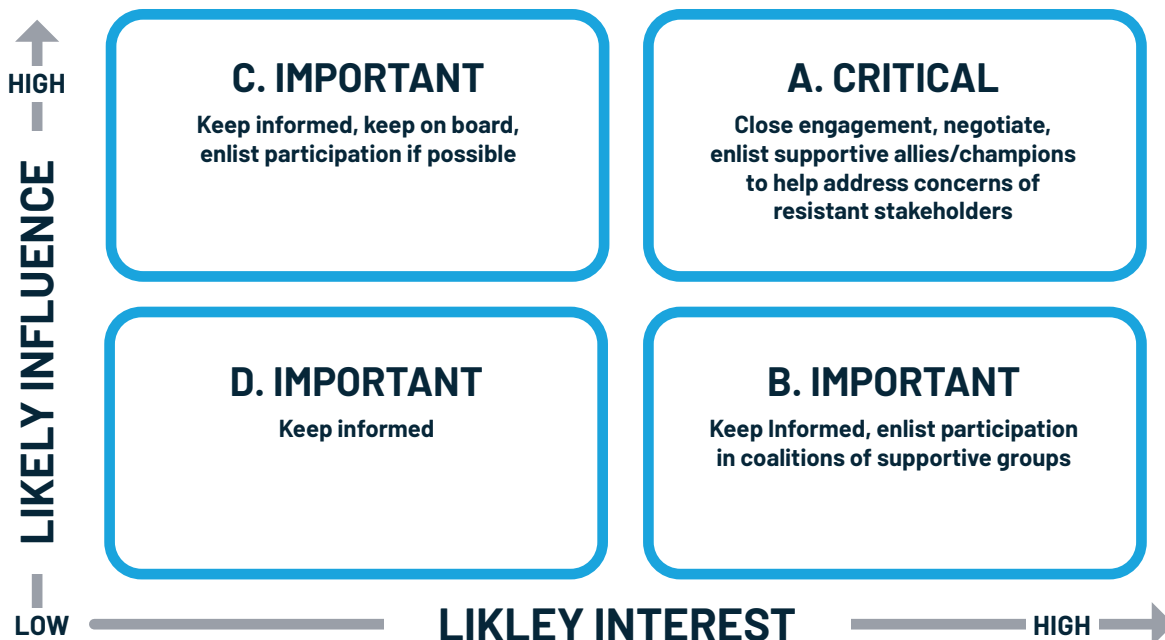
**Consider the possible range of interests—both positive and negative—of each group, such as:**

- Interest in improving service delivery, fostering research, or advancing policy goals
- History of exclusion or harm by government or agency officials
- Making the case for additional resources or identifying opportunities for savings
- Strengthening governmental administration, accountability, or efficiency
- Potential of being embarrassed about an unequitable or harmful system
- Discomfort or lack of experience with topic (civic data use or racial equity)
- Potential burdens of cooperation
- Inertia and organizational culture
- Privacy and security

When in doubt, reach out to potential stakeholders if you don't know or are unsure about their interests. Conversations of this kind will strengthen both decision making and relationships.

## 3. Prioritize your stakeholders.

Based on the information gleaned in the previous two steps, you will need to finalize the group of stakeholders who will be formally engaged. To do this, it is useful to think along two dimensions, influence and interest, as illustrated below. Powerful stakeholders with strong interests (quadrant A) demand the most attention.



(adapted from Bryson, J. 2004)

**Other tips for successful stakeholder mapping include:**

- **Finding early allies.** Those who are likely to be advocates for civic data sharing should be engaged early and encouraged to help address the concerns of other groups that may be influential but less supportive.
- **Champions.** When dealing with internal stakeholders, having dedicated champions can be an effective way to keep civic data sharing visible and relevant across several departments or organizations.
- **Transparency.** Determining which stakeholders actually participate in an engagement can become contentious, and so it is helpful to make the selection criteria as transparent as possible.
- **Group size.** Although diversity in stakeholder perspectives is important, an overly large group may be less effective in reaching consensus or navigating complicated discussion points. The size of the group engaged should reflect the breadth and sophistication of the participants, the complexity of the use case and engagement, and the consequence of data sharing activities.

## Gathering at the table

Once it is time to gather at the table, establish expectations and model a commitment to equity during all interpersonal engagement as well as structured activities and conversation. Providing guidance and setting that expectation for all participations early on will allow for helpful conflict resolution and reduce harm in the future.

**Here are some tips and further reading for modeling equity at your table:**

- Be aware of individual identities in the room and how they impact power dynamics.
  - Activities like the [Invisible Knapsack](#) or [Crossing the Line](#)<sup>32</sup> may help participants understand difference, and how identity and role shapes our understanding of the world.
- Recognize the difference between intent and impact. Both are important, and individuals have the ability to cause harm or hurt despite their best intention.
- Be prepared to address harm. Provide a plan to the group that defines harm, and outline processes for repair and restoring justice. Share these plans with the group, ask for feedback, and be willing to shift these practices over time.

**Some common practices used to address harm in racial equity-aware spaces include:**

- [Burn/aloe/recovery](#)<sup>33</sup>
- [Conflict transformation](#)<sup>34</sup>
- [Community reconciliation](#)<sup>35</sup>
- [Restorative justice](#)<sup>36</sup> (school-specific)

<sup>32</sup> Goldback, J. (2019)

<sup>33</sup> Lewis, T. et al. (2018)

<sup>34</sup> Racial Equity Tools (n.d.), [Conflict Transformation](#)

<sup>35</sup> Andrus, J. et al. (2001)

<sup>36</sup> Davis, F. (2019, April 21)



**Other useful resources for understanding accountability and repairing harm include:**

- [Tools for Addressing Chapter Conflict from Black Lives Matter](#)<sup>37</sup>
- [Challenging Neutrality, Examining Privilege and Encouraging Practitioner Self-Reflexivity: A Social Justice Approach to ADR](#)<sup>38</sup>

Celebrate community and the effort you are putting toward growth and change. A culture of appreciation is built when an organization takes time to appreciate the work and efforts of all involved and promotes sustainable, meaningful partnership over time.

**Here are other helpful resources to support this part of the process:**

- [White Supremacy Culture](#)<sup>39</sup> is a list of common ways that White Supremacy Culture shows up in organizational culture.
- For more on meeting facilitation, see page <sup>38</sup> of AISP and Future of Privacy Forum's [Nothing to Hide: Tools for Talking \(and Listening\) About Data Privacy for Integrated Data Systems](#).<sup>40</sup>
- For support in addressing implicit bias, racial anxiety, and stereotype threat, see this resource from King County, [Addressing Implicit Bias, Racial Anxiety, and Stereotype Threat Facilitator Guide](#).<sup>41</sup>
- For guidance on [how to set ground rules](#),<sup>42</sup> see this resource, *Guide for Setting Ground Rules*, from the Equity Literacy Institute.

<sup>37</sup> Hemphill, P. et al. (2018)

<sup>38</sup> Nagai-Rothe, E. (2010)

<sup>39</sup> Okun, T. (n.d.)

<sup>40</sup> Future of Privacy Forum & AISP (2018)

<sup>41</sup> King County Office of Equity and Social Justice (2015)

<sup>42</sup> Gorski, P. (n.d.)

## **Toolkit Activity 2: Mapping Assets and Engaging Community**

*For Toolkit Activity 2, we assume that your agency/organization/site/collaborative has purposefully and carefully identified a stakeholder group through completing Toolkit Activity 1 and has established a committed group of individuals that represent important perspectives on racial equity within your site context. These individuals should guide and support community engagement in service of centering racial equity in and across the data life cycle.*

There are many tools and approaches for collaboratively generating trust, but **a critical first step is developing shared understanding of the local racial, social, and historical context among stakeholders.**

To facilitate this process, we recommend using the [Community Engagement Framework: An Asset-based Approach](#), which was developed in 2015 by the Chief Education Office of Oregon (formerly the Oregon Education Investment Board). The nine-page community engagement framework centers the strengths of the community instead of starting from a needs-based approach. Strengths are then used to forge sustainable relationships between groups and individuals.

Below is an abbreviated version of the steps for following this framework to get your agency started on mapping assets and engaging community. We highly recommend downloading the [full document](#) and working through each step with your stakeholders.<sup>43</sup>

### **Step 1: Review the values and asset dimensions of the community engagement framework**

#### **Values:**

- **Integrity:** Acting with integrity means always being cognizant of power in exploitation of Black, Indigenous, people of color.
- **Transparency:** This framework strives to embrace potential conflicts, histories of actions/inaction, power dynamics, and the history of limited resources.
- **Collaboration:** This framework is a shift from seeking feedback on programs to authentic co-construction of ideas and plans based on assets, with the voices of Black, Indigenous, people of color centered.
- **Equity:** This framework encourages equity through an intentional examination of organizational practices in both a historical and sociocultural context.
- **Self-reflection:** This framework is meant to be a living document that will undergo multiple iterations as the work and community change.

<sup>43</sup> Oregon Education Investment Board, [Equity and Partnerships Subcommittee](#) (2015)

### Asset Dimensions:

- **Primary assets:** the structures and strengths of the agency/organization/site/collaborative
- **Secondary assets:** other strengths and structures that are leveraged regularly
- **Community assets:** existing relationships and connections in the community that might support goals
- **Historical and sociocultural assets:** the history and sociocultural context of the agency/organization/site/collaborative relevant to the work

**Step 2: Review the overall mission, vision, values, and goals of your organization and your data integration effort, using the worksheet below.**

ORGANIZATIONAL		Mission/vision, values, goals of the organization overall and in the context of the relationship
	OVERALL	RELATIONSHIP-SPECIFIC
MISSION / VISION	What is the overall mission/vision of the organization?	Is there a specific mission/vision for the relationship/project?
VALUES	What values structure the work of the organization?	How do these values manifest in this project?
GOALS	What are the stated and implied goals of the organization?	What are the goals specific to this relationship?

### Step 3: Identify the 6 P's in your community and their Asset Dimensions

#### The 6 P's:

- **People:** Individuals or groups who form the structures and strengths of the agency/ organization/site/collaborative
- **Place:** The geographic features of the land, physical spaces (offices, meetings spaces, locations, etc.), and climate that are primary assets for the agency/organization/site/ collaborative
- **Public:** The residents who stand to benefit
- **Promises:** Allocations of time or other resources and the outcomes to which the organization is accountable
- **Processes:** Theoretical frameworks and theories of action that guide the work
- **Programs:** Existing programs and projects that structure the work





PEOPLE	
PRIMARY	<p>Who are the people involved with the most direct influence on the project/relationship?</p>
SECONDARY	<p>Who else in the organization might support the work of the relationship? (i.e. research, finance, planning, graphic design, technology, other programs, etc.)</p>
COMMUNITY	<p>Who are some people in your community you already partner with or whose knowledge and expertise might support the relationship?</p>
HIST/SOCIO	<p>When are the people who started your organization?</p> <p>Are they still around?</p> <p>Who traditionally has been a part of designing and developing relationships?</p>

Step 4: Continue to revisit this framework regularly

## **Toolkit Activity 3: Identifying Root Causes through Factor Analysis**

*This activity requires drawing upon the community stakeholders identified in Activity 1, and follows the work of identifying the racial, social, and historical context of your site specific to data use in Activity 2.*

A common misstep during civic data use involves using data to describe problems without interrogating causal factors and root causes of the problem at hand. Data does not speak for itself, and it is important to carefully and collaboratively identify causal factors and root causes for observed outcomes in administrative data.

**causal factors:** conditions that contribute to an outcome.

If causal factors are not present, the outcome would be different.

**root cause:** primary factor that contributes to an outcome.

If the root cause was not present, the outcome would not occur.

When a social problem is observed, we often move quickly to develop solutions. For example, a community grappling with low third-grade reading scores may rush to implement an intervention such as academic-based summer programming. Factor analysis aims to uncover instances in which chosen solutions could be addressing a misunderstood problem and may therefore be ineffective. It does this by guiding stakeholders in asking why they might be observing the outcomes they do, or in other words: what is the story behind the data?

Factor analysis of systemic issues supports stakeholders in using data and lived experience to uncover the causal factors behind an issue so that solutions can go deeper.<sup>44</sup> For example, factor analysis could lead a community to address low third-grade reading scores by addressing the inequitable distribution of resources to support early childhood development in feeder neighborhoods before investing in downstream interventions.

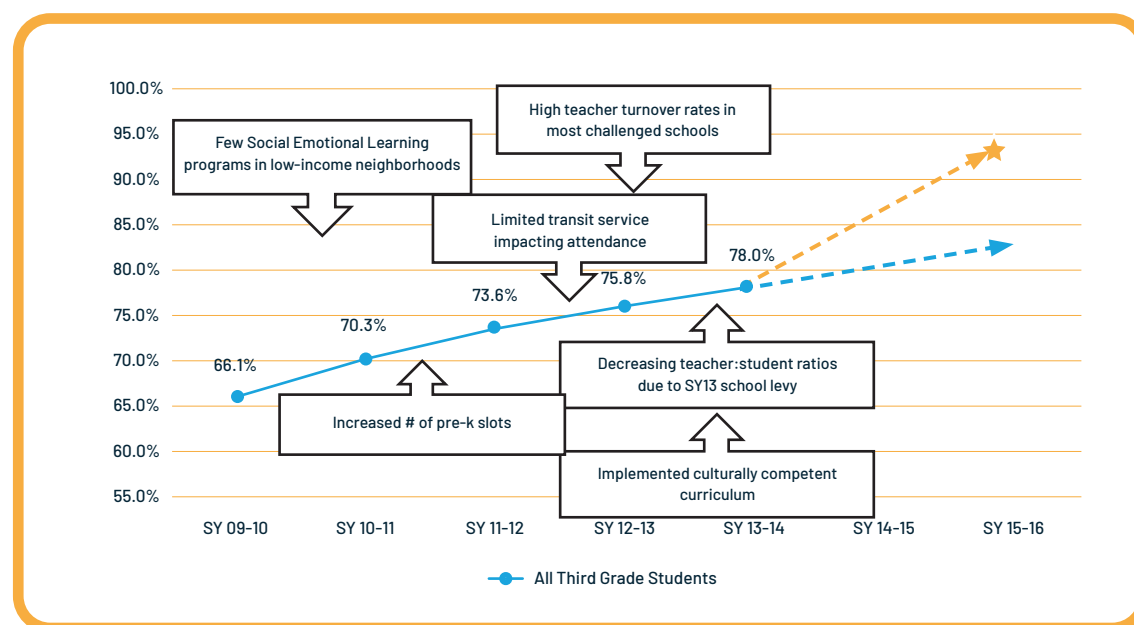
<sup>44</sup> This toolkit activity is based upon work of [StriveTogether](#) and [The Annie E. Casey Foundation](#), including [The Annie E. Casey Foundation \(2014\)](#) and [\(2018\)](#).

## Steps of Factor Analysis:

1. Define the current outcomes for a population and relevant subgroups of a population (e.g., race, ethnicity, gender, race × gender)
2. Identify causal factors. Collaboratively identify what is contributing to the outcomes.
  - **Gains being made:** What is contributing to the bright spot?
  - **No gains:** What is happening? Where is the population/subgroup losing ground?
3. Get to the underlying root causes. Ask “Why?” five times to understand the causal factors and the problem and solutions for the whole population or subgroup(s).
  - What is the underlying reason the problem or solution is occurring?
  - What is helping to shape the underlying reason?

Below are two examples where factor/root cause analysis<sup>45</sup> is helpful in explaining trends in outcomes over time. One explores the positive and negative factors influencing trends in third-grade literacy in a community, while the other explores policies, programs, and historical events impacting national trends in the number of people experiencing homelessness.

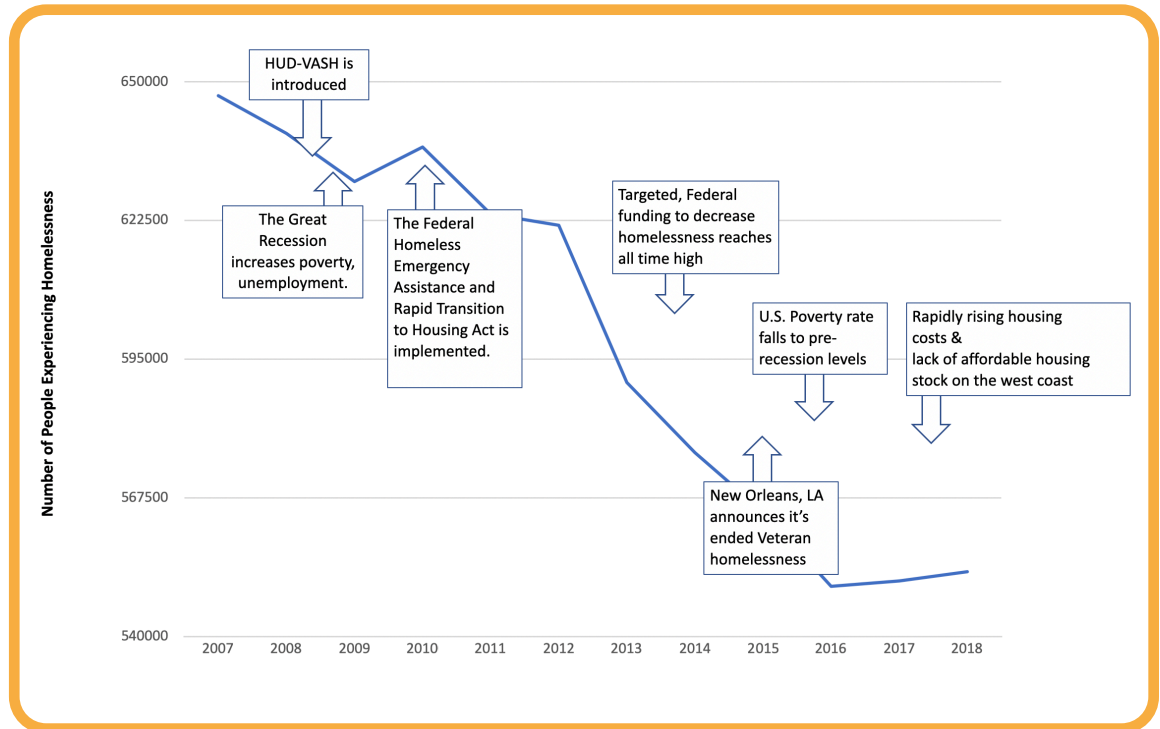
## EXAMPLE 3<sup>RD</sup> GRADE LITERACY



(Figure used with permission, StriveTogether, n.d.).

<sup>45</sup> Note that we are not referring to factor analysis as a statistical method in this context.

## HOMELESSNESS IN THE U.S. FROM 2007–2018:



(Figure used with permission, TC Burnett, AISP, 2020).



## References

Anderson, D., Bonaguru, J., McKinney, M., Nicklin, A., Wiseman, J. (2018). *Ethics & Algorithms Toolkit*. Center for Government Excellence, Johns Hopkins University.

Andrews, K., Parekh, J., Peckoo, S. (2019). *A guide to incorporating a racial and ethnic equity perspective throughout the research process*. Child Trends.

The Annie E. Casey Foundation. (2018). *Race Equity Crosswalk Tool*, Annie E. Casey Foundation.

The Annie E. Casey Foundation. (2014). *The Race Equity and Inclusion Action Guide*.

Andrus, J., K. Downes, M. Umbreit. (2001). *Community Reconciliation Through Facilitated Dialogue & Restorative Justice*. Andrus Family Fund Board Exploration Triad, Working Paper.

Barocas, S., & Selbst, A. D. (2016). *Big data's disparate impact*. *Calif. L. Rev.*, 104, 671.

Benjamin, R. (2019). *Race After Technology: Abolitionist Tools For the New Jim Code*. Wiley.

Berkeley Underground Scholars. (2019). *Language Guide for Communicating About Those Involved in the Carceral System*. Underground Scholars Initiative.

Chicago Beyond. (2019). *Why am I always being researched? A guidebook for community organizations, researchers, and funders to help us get from insufficient understanding to more authentic truth*. Chicago Beyond Equity Series, Vol. 1.

Davis, F. (2019, April 21). *How schools are using restorative justice to remedy racial disparities in discipline*. Salon. Excerpted from Davis, F. (2019). *The Little Book of Race and Restorative Justice: Black Lives, Healing, and US Social Transformation*. Good Books.

Diakopoulos, N., Friedler, S., Arenas, M., Barocas, S., Hay, M., Howe, B., Jagadish, H. V., Unsworth, K., Sahuguet, A., Venkatasubramanian, S., Wilson, C., Yu, C., & Zevenbergen, B. (2017). *Principles for accountable algorithms and a social impact statement for algorithms*. FAT/ML.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Future of Privacy Forum & Actionable Intelligence for Social Policy. (2018). *Nothing to Hide: Tools for Talking (and Listening) About Data Privacy for Integrated Data Systems*.

Goldbach, J. (2019). *The MSW@USC Diversity Toolkit: A Guide to Discussion Identity Power and Privilege*. University of Southern California, Suzanne Dworak-Peck School of Social Work.

Good, M. & Lucadamo, S. (2018). *Moving a Child Welfare System to be More Affirming of the LGBTQ Community: Strategies, Challenges, and Lessons Learned, Initiatives from 2009 through 2016*. Allegheny County Department of Human Services.

Gorski, P. (n.d.). *Guide for Setting Ground Rules*. Awareness Activities. EdChange.

- Government Alliance on Race & Equity. (2018, May). [GARE Communications Guide](#). Local and Regional Government Alliance on Race & Equity.
- Guobadia, O. (2018). [Kimberlé Crenshaw and Lady Phyll talk intersectionality, solidarity, and self-care](#). them.
- Hawn Nelson, A., Jenkins, D., Zanti, S., Katz, M., Burnett, T., Culhane, D., Barghaus, K., et al. (2020). [Introduction to Data Sharing and Integration](#). Actionable Intelligence for Social Policy. University of Pennsylvania.
- Hayes-Greene, D., & Love, B. P. (2018). [The Groundwater approach: Building a practical understanding of structural racism](#). Racial Equity Institute.
- Hemphill, P. et al. (2018). [Tools for Addressing Chapter Conflict](#). Black Lives Matter.
- Hetey, R., & Eberhardt, J. (2018). [The numbers don't speak for themselves: Racial Disparities and the persistence of inequality in the criminal justice system](#). *Current Directions in Psychological Science*, 27(3), 183–187. <https://doi.org/10.1177/0963721418763931>
- Holland, S., Hosny, A., Newman, S., Joseph, J., Chmielinski, K. (2018). [The dataset nutrition label: A framework to drive higher data quality standards](#). MIT Media Lab and the Berkman Klein Center.
- Hyams, K., Prater, N., Rohovit, J., Meyer-Kalos, P.S. (2018). [Person-centered language](#). Clinical Tip No. 8 (April, 2018): Center for Practice Transformation, University of Minnesota.
- Katell, M., M. Young, D. Dailey, B. Herman, V. Guetler, A. Tam, C. Binz, D. Raz, P.M. Krafft. (2020). [Toward Situated Interventions for Algorithmic Equity: Lessons from the Field](#), *Association for Computing Machinery*. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. <https://dl.acm.org/doi/abs/10.1145/3351095.3372874>
- King County Office of Equity and Social Justice. (2015). [Inclusion, Belonging, and Excellence for One King County: Addressing Implicit Bias, Racial Anxiety, and Stereotype Threat](#). Facilitator Guide.
- Kingsley, C. (2017). [First, do no harm: Ethical guidelines for applying predictive tools within human services](#). MetroLab Network.
- Kirkland, D. (2019). [No Small Matters: Reimagining the Use of Research Evidence From A Racial Justice Perspective](#). William T. Grant Foundation.
- The Leadership Conference on Civil & Human Rights. (2018). [Shared statement of civil rights concerns for the use of pretrial "risk assessment" instruments](#).
- Lewis, T., Gangadharan, S. P., Saba, M., Petty, T. (2018). [Digital defense playbook: Community power tools for reclaiming data](#). Detroit: Our Data Bodies.
- Lief, L. (2020). [How Philanthropy Can Help Lead on Data Justice](#). Stanford Social Innovation Review.
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019, January). [Model cards for model reporting](#). In Proceedings of the conference on fairness, accountability, and transparency (pp. 220-229).

- Nagai-Rothe, E. (2010). *Challenging Neutrality, Examining Privilege, and Encouraging Practitioner Self-Reflexivity: A Social Justice Approach to ADR*. Mediate.
- National Center for Health Marketing. (2007). *Plain Language Thesaurus for Health Communications*.
- Nelson, J., & Brooks, L. (2015). *Racial Equity Toolkit: An opportunity to operationalize equity*. Local and Regional Government Alliance on Race & Equity.
- Nelson, J., Spokane, L., Ross, L., & Deng, N. (2015). *Advancing Racial Equity and Transforming Government: A Resource Guide to Put Ideas into Action*. Local and Regional Government Alliance on Race & Equity.
- Okun, T. (n.d.) *White Supremacy Culture*. changeworkDR.
- Plain Language Action and Information Network. (n.d.) *Checklist for plain language*.
- Oregon Education Investment Board, Equity and Partnerships Subcommittee. (2015). *Community Engagement Framework: Equity, Transparency, Collaboration, Integrity, Self-reflection*.
- ProInspire: Equity in the Center. (2019). *Awake to Woke to Work: Building a Race Equity Culture*.
- Racial Equity Tools (n.d.). *Core concepts: Racism & Conflict Transformation*.
- Rainie, L., Keeter, S., & Perrin, A. (2019, December 31). *Americans' trust in government, each other, leaders*. Pew Research Center.
- Readability Formulas. (n.d). Automatic readability checker. [Readabilityformulas.com](https://readabilityformulas.com).
- Reisman, D., Schultz, J., Crawford, K., & Whittaker, M. (2018). *Algorithmic impact assessments: A practical framework for public agency accountability*. AINow.
- Richardson, R. ed. (2019). *Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force*. AI Now Institute. [https:// ainowinstitute.org/ads-shadowreport-2019.html](https://ainowinstitute.org/ads-shadowreport-2019.html)
- Staats, C., Capatosto, K., Tenney, L., & Mamo, S. (2017). *State of the science: Implicit bias review 2017 edition*. Kirwan Institute, The Ohio State University.
- Steif, K. & Goldstein, S. (2019). *Algorithmic fairness: A code-based primer for public-sector data scientists*. Urban Spatial.
- Stevenson, B. (2014). *Just mercy: A story of justice and redemption*. Spiegel & Grau.
- United Kingdom Department for Digital, Culture, Media & Sport. (2018). *Data Ethics Workbook*.
- Wilber, S. (2013). *Guidelines for Managing Information Related to the Sexual Orientation and Gender Identity and Expression of Children in Child Welfare Systems*. Putting Pride Into Practice Project, Family Builders by Adoption, Oakland, CA.

## Appendix I: Work in Action Throughout the Data Life Cycle

The following section includes site-based examples of Work in Action throughout the data life cycle, expanding upon the briefs included within the main part of the Toolkit.

### Racial Equity in Planning: WORK IN ACTION

#### Broward Data Collaborative by Sue Gallagher

**Who:** Broward Data Collaborative, Children's Services Council of Broward County

**Where:** Broward County, Florida

**Organization Type:** Government Agency, Community-Based Organization

**Domains:** Child welfare, behavioral health, juvenile justice, early learning, school, human services, prevention programs.

**Goal/Impact:** Through equitable collective action, diverse research strategies, and the use of high-quality data, to generate actionable intelligence that improves quality-of-life outcomes and community conditions.

The Broward Data Collaborative (BDC) has worked to provide a seat at the proverbial (and literal) table for residents whose data is represented in the child-serving integrated data system (IDS). The BDC's recognition that their county-level systems produce racially disparate outcomes has guided their work to begin building nonhierarchical relationships between the community and professionals, inviting system participants to tell strengths-based stories about themselves and their community and use data to co-create system and policy improvements.

The BDC was created in 2017 with support from AISP through the AISP Learning Community initiative. The BDC seeks to improve the outcomes of residents in Broward County by integrating child-serving data from child welfare, behavioral health, juvenile justice, schools, early learning systems, county human services, and prevention programs. Data uses for the BDC include evaluation, research, and care coordination. The Children's Services Council of Broward County, the BDC's backbone organization, is an independent special taxing district that funds prevention programs for children and families in Broward County and has led the Broward Children's Strategic Plan for 20 years. Currently, the BDC's partners include senior leadership, researchers, and technology professionals from the human services agencies providing data to the BDC.

The BDC has rejected a race-neutral or colorblind design for data access and use. Instead, the BDC is *intentionally* designing an IDS framework that accounts for the historical and ongoing structural racism in the community, human services system, and service organizations. Too often, data infrastructures exist as racialized hierarchies with predominantly White researchers/policy makers and system professionals yielding power over participants represented in the data who are disproportionately BIPOC. Additionally, there is typically social and spatial segregation between users of the IDS, the system professionals, and people

whose data are in the system. Without proximity to system participants' lived experiences, system analysts and decision makers are at great risk of perpetuating stereotypes and bias through data-informed decision making that is not relational, holistic, or compassionate.<sup>46</sup> For example, individual behaviors or choices look very different from a deficit-based perspective (e.g., a youth in foster care did not go to college) vs. a strengths-based and contextual perspective that includes the lived experience of the youth (e.g., a youth did not go to college because they needed to take care of their younger siblings in order to reunite the family).

Based on the BDC's two core values, equity and transparency, the BDC is prototyping racial equity structures for an IDS by employing Community Participatory Action Research (CPAR). The CPAR framework builds the capacity of system participants to be effectively involved in the governance, research, and evaluation of an IDS in order to address the racialized hierarchy between professionals and service recipients. One of the BDC's desired outcomes is for the people whose data are in the IDS (e.g., youth aging out of the child welfare system, youth in the juvenile justice system) to be integrally involved in the use, interpretation, and evaluation of their data. The BDC is working to ensure that these groups have a seat at the decision-making table, with the resources and capacity-building necessary to fully participate. Recognizing that every child-serving system (e.g., child welfare, juvenile justice, education) produces racially disparate outcomes, the CPAR work is supported by training partners on the history and structures of racism (both nationally and locally) to build a common language and framework for system participants and system professionals.

In Broward County, Black children are more likely to be removed from their parents than White or Latino/a children. Furthermore, the highest removal rates of Black children occur in the historically Black communities that are concentrated in four zip codes established under the Jim Crow segregation laws of the 1930s. To address this systemic racism, the BDC's CPAR work engaged youth and parents who participated in the child welfare system. With funding from the Florida Institute for Child Welfare, system participants were trained as CPAR co-researchers alongside system professionals (i.e., staff investigators who determine whether children will be removed from the home, child welfare services staff, and researchers). They cultivated the leadership, advocacy, and research skills needed to leverage their expertise, social networks, and passion to improve the outcomes and processes of the child welfare system. Simultaneously, CPAR allowed system professionals and researchers to build inspiring relationships, learn from those with lived experience and expertise, and see the people they serve as more than individual data points. To date, the CPAR initiative continues to use these strategies to build nonhierarchical relationships, overcome professional/participant segregation, and tell a strengths-based story that is generating policy and system improvements.

## Tacoma Equity Index by Alison Beason

**Who:** City of Tacoma

**Where:** Tacoma, Washington

**Organization Type:** Government Agency

**Domains:** Livability, Accessibility, Economy, Education, Strategic Plan

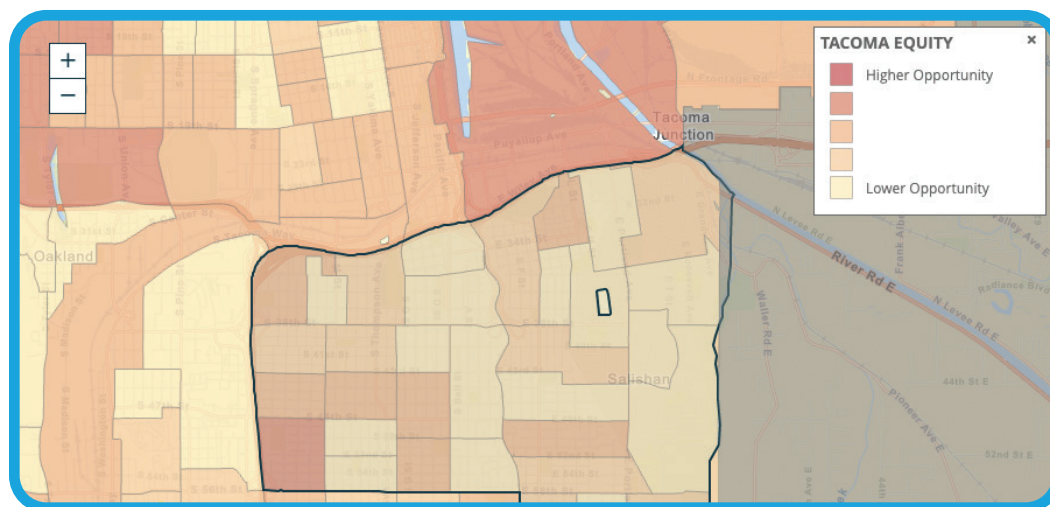
**Goal/Impact:** To design an Equity Index to support data-driven policy decisions and create a more equitable and inclusive Tacoma where all people have the opportunity to succeed.

In 2017, the City of Tacoma elected Mayor Victoria Woodards to office. The leadership of Mayor Woodards and her predecessor, Marilyn Strickland, both Black females, has led to strides in celebrating Tacoma's history of diversity, while also recognizing Tacoma's legacy of adversity. This is evident in the development of the City of Tacoma's 2025 Strategic Visioning Framework, which provides an overarching direction for the city over the next decade, with the goal of increasing accessibility, education, economy, and livability.

<sup>46</sup> Stevenson, B. (2015); Staats, C. et al. (2017)

The strategic plan was created using a collaborative and inclusive approach, including resident surveys. It centers on themes of equity, opportunity, partnerships, and accountability, and specifically reflects community members' desire for racial fairness and social justice across all public programs. Over 71 percent of Tacoma residents surveyed during the planning process believed that equitable access to facilities, services, and infrastructure was "essential" or "very important." With this charge, the City of Tacoma sought to improve service delivery to make Tacoma a more inclusive and accessible place to live, work, play, and visit for all races, ethnicities, abilities, and levels of income. Through these efforts, the City of Tacoma's Equity and Empowerment Initiative was established, including the Office of Equity and Human Rights. The city partnered with organizations in the region to examine "health equity"—the opportunity for all residents to reach their full health potential. The city recognized that income, education, housing, transportation, and other factors can create opportunities or barriers, and sought to better understand the relationship among these systems.

Building on this work, the city then collaborated with the Ohio State University's Kirwan Institute for the Study of Race and Ethnicity to create the Equity Index. The Equity Index, drawing from the concept of opportunity mapping, is an interactive tool that uses administrative data to highlight successes and obstacles to mobility in different areas of Tacoma. The Index includes 20 indicators<sup>47</sup> that are strongly correlated to equitable outcomes and aligned with the Tacoma 2025 Strategic Goals: accessibility, education, economy, and livability. The tool helps facilitate data-driven decision-making processes that enable leaders to better distribute resources and plan program funding that minimizes inequities and maximizes opportunities.



Screenshot from [Tacoma Equity Map](#)

Each census block group is scored on its access to opportunity, on a range from "Very Low" to "Very High" Equity. "Very High Equity" indicates locations that have access to better opportunities and where residents are more likely to succeed. The data indicators for this designation include high-performing schools, a safe environment, access to adequate transportation, safe neighborhoods, and sustainable employment. In contrast, "Low Equity" areas exhibit more obstacles and barriers within each of these areas, and communities designated as "Low Equity" have limited access to institutional or societal investments, which limits their quality of life.

Tacoma was mindful of not reinforcing historical stereotypes of low-income communities or communities of color when constructing the Equity Index. To counteract harmful narratives, the index shifts its use of

<sup>47</sup> To view all 20 indicators, see <http://cms.cityoftacoma.org/OEHR/EquityIndex/EquityIndexDataDefinitions.pdf>



shading and color. The Index does not use green, which has generally positive associations, to refer to higher-opportunity and predominantly White communities. Instead, red is used to visualize areas of high opportunity and thereby call out a need for increased resources. The use of colors was a simple way to confront the public mindset that neighborhoods with less opportunity are “bad” and high-opportunity neighborhoods are “good.”

As a result, the tool has helped educate and change law makers’ and community groups’ approaches to decision making. It has also led to the creation of consistent policies and processes while highlighting key components of equitable policy making. Further, it provides accurate data that helps inform the conversation about equity and disparities. Through implementing this tool, policy makers will become more aware of the importance of consulting with affected groups and staff, while giving organizations a data-informed tool to apply an equity lens to their day-to-day operations. In this way, the Equity Index supports the actualization of the City of Tacoma’s Strategic Visioning Framework.

## Racial Equity in Data Collection: WORK IN ACTION

### Mecklenburg County Community Support Services by Courtney LaCaria & Mary Ann Priester

**Who:** Mecklenburg County Community Support Services

**Where:** Mecklenburg County, North Carolina

**Organization Type:** Government Agency

**Domains:** Housing, homelessness, social services

**Goal/Impact:** To collect high-quality, useful, and timely data and use these data to inform funding, programs, and policy.

In 2014, Mecklenburg County Community Support Services stepped up to fill a gap in the community by investing in two positions dedicated to improving housing and homelessness data. The Homeless Management Information System (HMIS) Administrator is primarily responsible for ensuring that the community has high-quality, accurate data on housing and homelessness, and the Homelessness & Housing Research Coordinator position is focused on connecting that data to stakeholders in the community to use it.

While responsible for specific areas of work related to housing instability, these positions are collectively focused on ensuring that Charlotte-Mecklenburg collects high-quality, useful, and timely data, *and* that the data are connected to the community in order to better inform funding, programs, and policy.

The HMIS Administrator guides county staff and community partners in expanding participation within HMIS; improving data collection processes, data quality, and data analysis; and implementing best practices and new HMIS features. This role supports Mecklenburg County’s priority to end homelessness by ensuring that high-quality data and analysis of service strategies are available to the county and its partner agencies. In addition, this work supports the annual \$4 million federal HUD funding for homeless services and guides resource allocation decisions. **Essential functions include:**

- Ensuring data accuracy, data quality, and compliance with HUD reporting measures for Continuum of Care
- Building partner agency capacity to collect and utilize data for continuous quality improvement
- Ensuring that the HMIS program is maintained and used so that community programs and community needs can be measured and evaluated

- Reporting outcomes to project and community partners
- Recommending changes to the homeless service system based on outcomes
- Serving as point of contact for inquiries and issues from end-users related to HMIS
- Providing technical end-user support for software and ensuring quality, accessibility, and functionality of HMIS for partner agencies

The Housing & Homelessness Research Coordinator connects research and data with community stakeholders to inform programming, policy, and funding decisions related to addressing housing instability and homelessness.

**Essential functions include:**

- Designing strategies to connect data with community stakeholders for policy-making, funding decisions, and programmatic change
- Disseminating community reports and toolkits to inform decision making and strengthen community performance
- Managing a community online hub for local data and research on housing and homelessness: Charlotte-Mecklenburg Housing & Homeless Dashboard ([MecklenburgHousingData.org](https://mecklenburghousingdata.org))
- Leading presentations and workshops that synthesize and translate data and research into relevant, useful, and actionable steps
- Serving as project manager for the Point-in-Time Count, bringing together community volunteers and providers to plan and implement a count of persons experiencing homelessness in Mecklenburg County on one night in January
- Providing analysis, interpretation, and visualization to synthesize and present actionable data and research
- Providing technical assistance to agencies to strengthen outcome measurement and program evaluation systems

Together, the HMIS Administrator and the Housing & Homelessness Research Coordinator have improved Charlotte-Mecklenburg's housing and homelessness data collection practices. The investment has led to important disaggregation across multiple intersectional categories, describing disparate impact of housing instability on Black residents—specifically, Black women with children—therefore informing more equitable resource allocation.<sup>48</sup> Moreover, this work has connected community stakeholders to the data and allowed for increased data-driven decisions around funding, programs, and policy that impact community members experiencing homelessness.

**Allegheny County Department of Human Services Office of Equity & Inclusion by Shauna Lucadamo, Jessica Ruffin, Ellen Kitzerow, Julia Reuben, & Shannon Flynn**

**Who:** Allegheny County Department of Human Services

**Where:** Allegheny County, Pennsylvania

**Organization Type:** Government Agency

**Domains:** Child welfare; LGBTQ+ children, youth, and families

**Goal/Impact:** To support provision of culturally responsive and affirming services to LGBTQ communities

<sup>48</sup> See <https://mecklenburghousingdata.org/>



In 2009, Allegheny County's Department of Human Services (DHS) began their work to collect sexual orientation, gender identity, and gender expression (SOGIE) data from families involved with child welfare. The process was part of a department-wide effort to provide culturally responsive and affirming services and supports to LGBTQ communities involved with DHS.<sup>49</sup> At the time, gender and sexuality were not openly discussed as part of DHS practice, and the experiences and needs of LGBTQ communities were not well understood. Collecting SOGIE data was just one of several pieces aimed at changing this.

The effort started by creating a data collection subcommittee, which ultimately recommended that DHS update case management systems to start collecting SOGIE data and provide training and support to caseworkers on how to incorporate these changes into practice. The original hope was to update the case management system and begin collecting data in the third year of the work, but it became clear early on that changes would not come so quickly. DHS had to address several challenges, such as concerns that a youth's SOGIE information would be mishandled, the implications of SOGIE data being shared with external stakeholders (e.g., parent and child advocates, out-of-home placement providers), and the complexity and cost associated with updating IT systems. As DHS engaged with stakeholders to understand their concerns, they realized that aligning systems with desired practice was more complex than simply changing drop-down values in an IT system.

As part of this process, DHS focused on key questions to address these challenges and guide the development of SOGIE data collection work:

- Who is informing the changes being made in the information system?
- Who has access to the information system?
- Will functionality be built-in to mark information as confidential if necessary?
- What guidance is provided to workers?
- How do you manage safety risks related to outing the youth?
- How will workers be held accountable for the proper use of documented information?

In 2013, DHS partnered with the Center for the Study of Social Policy to pilot guidelines for managing information related to gender and sexuality through an initiative called “getREAL.”<sup>50</sup> This initiative employed a full-time project manager dedicated to improving practice and building stronger relationships throughout the LGBTQ communities.

This partnership supported DHS in making the necessary IT system changes to collect SOGIE data. Throughout this process a valuable lesson emerged, one that continues to challenge the DHS team: the importance of working with the individuals responsible for building and maintaining IT infrastructure so that they understand why inclusive changes are important and how to appropriately design the fields. The role of IT staff is critical, as IT changes rapidly and program staff may not be constantly engaged in all IT design projects. For changes like this to succeed, it is essential to go beyond handing staff a list of drop-down field values, and instead help them understand why the current recommendations are what they are, and how and when to seek additional consultation.

Recognizing that information was a critical piece of effectively improving practice, the county field tested the Guidelines for Managing Information Related to the Sexual Orientation and Gender Identity and Expression of Children in Child Welfare Systems.<sup>51</sup> The field testing was intended to show the broader child welfare field what it takes to shift practice toward regularly gathering information related to SOGIE and how using the information appropriately can improve practice and outcomes. Now, years following implementation, collection and use of these data remain a challenge for the agency and its staff. The key questions remain guideposts for the work, and the information provided by the collected data (or lack thereof) highlights the necessity of those questions and provides context and support for properly addressing those questions.

49 Good, M. & Lucadamo, S. (2018)

50 Ibid.

51 Wilber, S. (2013)

Technical challenges involving collection timing and additional challenges around staff's confidence with asking questions about SOGIE and recording the information have resulted in low rates of data collection. Improvements to the data system, increased reporting of missing data to staff and leadership, and interventions and consultants available on and off-site are among the current efforts to improve our ability to identify and improve the experience of LGBTQ youth involved in child welfare.

## Racial Equity in Data Access: WORK IN ACTION

### KYSTATS by AISP with contributions from Jessica Cunningham

**Who:** Kentucky Center for Statistics, KYSTATS

**Where:** State of Kentucky

**Organization Type:** Government Agency

**Domains:** Early childhood, K-12 and postsecondary education, employment and training, vocational rehabilitation, healthcare and educator licensure

**Goal/Impact:** To collect and link data to evaluate education and workforce efforts in the Commonwealth so policy makers, practitioners, and the public can make better-informed decisions.

The Kentucky Center for Statistics (KYSTATS) was created in 2012 to expand upon the work of the Kentucky P-20 Data Collaborative, including maintaining the Kentucky Longitudinal Data System, a statewide longitudinal data system that facilitates the integration of workforce data and data from K-12 and postsecondary education systems. KYSTATS has the authority to collect and link data in order to evaluate education and workforce efforts in the Commonwealth. This includes developing reports and providing statistical data about these efforts so policy makers, agencies, and the general public can make better-informed decisions about Kentucky's education systems and training programs. While collecting and researching this information does not give Kentucky a "crystal ball" vision of the future, it does provide the state with a map connecting past outcomes with future goals.

KYSTATS strives for transparency in their data request process in order to more easily connect users with the rich datasets they maintain. A clear [data access and use policy](#)<sup>52</sup> is posted online, where users can find information for accessing data based on their user type (i.e., agency staff, the public, or individuals) as well as an overview of data collected and maintained by KYSTATS. In addition, the organization maintains a detailed flow chart of steps for filling data requests from internal and external parties, including the review process and timeline. Separate request forms for individual- and aggregate-level data allow users to clearly see the information required in order to [request data](#) and to submit their requests online.

Fees for data requests are based on a cost-recovery model and are therefore within budget for most users. KYSTATS also maintains a data dictionary on their website, which includes detailed, high-quality metadata to inform requestors about available data no matter their level of data literacy. Those interested in monthly employment data releases will find a clear and updated press release schedule on the KYSTATS website. This allows users to know when to expect new data to become available and to view a historical log of Kentucky labor market data releases. All of this information is posted prominently on the KYSTATS website rather than buried beneath numerous clicks or jargon-filled documents. This type of transparency in process is an important element of promoting equity, as it helps all types of users understand what data are collected and how and where they can be accessed, underscoring that these data are not available only to elite researchers or those with special connections to the agency.

<sup>52</sup> Kentucky Center for Education and Workforce Statistics (2017)

Moreover, KYSTATS' transparency in data access processes extends to how they handle the de-identification and anonymization of data. Upholding the privacy of individuals represented in the data is foundational to racial equity, as the improper release of these data has the potential to exploit communities who have endured prior system-level harm and are not often represented by a seat at the table. Thus, KYSTATS ensures that education and workforce information is collected with the utmost care for security and privacy and follows strict reporting protocol. Prior to release, data are stripped of all identifiers such as the person's name, date of birth, social security number, and the agency or institution identification. Then, the data are given a unique identifier that protects the person's and institution's information. Once the data are stripped of identifiers, they are placed in a de-identified warehouse where analyses, reports, requests, and evaluations are created.

The focus of KYSTATS' research is on the numbers, not the individuals. Personally Identifiable Information (PII) is used to initially match people within the system and is then replaced by unique system identifiers. No PII moves with the data to the De-Identified Reporting System. Rather, reports and analyses are generated from data that no longer contains any PII. The transparency of the process and preservation of privacy distinguishes KYSTATS' work toward creating equity in data access.

### Allegheny County Department of Human Services, Office of Analytics, Technology and Planning by AISP with contributions from Samantha Looney, Brian Bell, & Jamaal Davis

**Location:** Allegheny County, Pennsylvania

**Organization Types Represented:** Government Agency

**Domains Represented:** Child welfare, family support, housing/income support, behavioral health, criminal justice, aging and intellectual disabilities

**Goal:** To provide clients and providers access to the robust data held by the county's Department of Human Services, allowing clients to understand what data is collected and providers to give more holistic care.

Allegheny County's Client Experience Analytics Unit is developing a new client portal that will allow Department of Human Services (DHS) staff, network providers, and clients to access data housed in the county's data warehouse.<sup>53</sup> The client portal will contain data from local school districts, family support programs, housing, income support programs, child welfare agencies, drug and alcohol treatment, the criminal justice system, mental health providers, and more. Moreover, the new system will provide transparency and greater data access to clients and providers, increasing the capability for both to coordinate care.

Allegheny County DHS has consulted clients throughout the design process, making use of individual interviews and client feedback on prototypes. Centering the community and seeking feedback from future portal users is a priority for the Client Experience Analytics Unit, as it ensures that this new interface reflects the resources most useful to those served.

Allegheny County recognizes community members' right to their own information and is building a system that will allow residents to see the information collected about them and greater agency to ensure the information is correct. The portal will serve as a mechanism for clients to provide feedback on the services they receive—providing a needed check on government programs and services and giving residents an outlet to comment on the quality of services they receive. The client portal will bring client information into one central database, allowing clients to play a bigger role in their own care coordination. The client portal is built on transparency, both of process and of information, to allow community members full access to their own data.

<sup>53</sup> See [AISP, Network Sites, Allegheny County, PA](#)

## Birth through Eight Strategy for Tulsa (BEST) / Asemio Privacy-preserving Record Linkage Pilot by Jessica England & Dan Sterba

**Location:** Tulsa, Oklahoma

**Organization Types Represented:** Government Agency, Community-Based Organizations, Private Sector, Philanthropy

**Domains Represented:** Public health, early childhood education, K-12 education, basic needs, youth services

**Goal:** To apply innovation in data analytics with the aim of decreasing cost, increasing speed, and improving security and ethical protections for individuals served as well as data contributors.

The Birth through Eight Strategy for Tulsa (BEST) / Asemio Privacy-preserving Record Linkage Pilot utilized a data sharing and analytics platform to address common challenges for data sharing across entities with disparate organizational boundaries and political, legal, and ethical concerns. Specifically, the platform addresses these concerns by de-identifying personally identifiable information (PII) at the source using NIST-approved cryptographic security standards, and then computing on the resulting encrypted data by utilizing secure multi-party computation.

In 2019, social service leaders in the Tulsa community came together to identify and then answer community-level questions regarding race, equity, and service overlap. Eight organizations representing 32 programs contributed data from a 12-year period for this effort. Organization types represented included county/local government, nonprofit agencies, private sector, and philanthropy. The included data sets are protected by HIPAA, FERPA, 42 CFR Part 2, and other federal, state, and local regulations. Programs serving community members span public health, early childhood education, K-12, basic needs, and youth services. The combined dataset included over 220,000 unique individuals and answered five community questions that centered on program sequencing, outcome achievement, and racial and gender equity:

1. What is the unique number of individuals served (i.e., total population) and the overlap between each pair of data contributors for all time?
2. What is the unique number of individuals served (i.e., total population) and the overlap between each pair of data contributors for 2018?
3. What are the differences in program completion rates for different racial, ethnic, and gender groups for the total population of individuals served in 2018?
4. What are the differences in outcome achievement rates for different racial, ethnic, and gender groups for the total population of individuals served in 2018?
5. How many individuals previously served by two of the participating early education data contributors were subsequently served by two of the participating youth and adolescent data contributors?

The privacy-by-design infrastructure and approach enabled the answering of these important community questions by linking data across contributors without ever sharing PII. The data upload process lasted, on average, two hours per agency. Pre-processing was performed using a proprietary tool designed to clean, transform, de-identify, and load data without the need for sharing any PII. This approach to community analytics (i.e., privacy-preserving record linkage) improves the security, ease, and return on investment of collaborative action and has been successfully deployed in two prior pilots, paving the way for this third project.

The novel ability to complete analysis without sharing PII was possible through the combination of cryptographic and record-linking technology. Remarkably, the application of a trust-centered process in

combination with the platform's improved ethical and security protections produced results within two months of the first stakeholder meeting.

This project demonstrates a new, effective way to collaborate and share data between community partners. It was driven by agency leaders, with no requirement to participate and the primary incentive being access to information that could be used for improved service delivery. Funding was provided by a local philanthropic partner who transferred decision-making power to the data collaborative leadership while still providing access to critical philanthropic support. A social enterprise served as a neutral community convener and provided the technology platform as well as expertise in the facilitation of trust, governance, and data analytics. This project demonstrated a new platform and approach that more quickly integrates data, at a lower cost, and in a manner that enhances the privacy and security protection for both the individuals and the organizations who serve them.

## Racial Equity in Algorithms / Use of Statistical Tools: WORK IN ACTION

### Allegheny County Child Welfare Algorithm by AISP with contributions from Katy Collins

**Location:** Allegheny County, Pennsylvania

**Organization Types Represented:** Government Agency

**Domains Represented:** Criminal justice system, behavioral health, public assistance, child welfare, education

**Goal:** To use predictive analytics to allow caseworkers to engage in data-driven decision making that creates a more equitable, efficient, and successful child welfare system.

In 2016, Allegheny County, Pennsylvania's Department of Human Services (DHS), in partnership with researchers from Auckland University (New Zealand) and the University of Southern California, created the Allegheny Family Screening Tool (AFST) to support the decision-making process of county child welfare hotline screening staff when assessing child maltreatment allegations. The creation of the AFST was the county's response to a child welfare system plagued by racial and economic disproportionality, intrusive investigations, and a string of maltreatment-related deaths. The AFST harnesses a selection of the data stored in the Allegheny County Data Warehouse, an integrated data system comprising data from more than 20 sources, to provide caseworkers with a risk assessment score based on a family's previous contact with public services and systems. This assessment serves as a more objective measure of risk, allowing call screeners to engage in data-driven decision making.

In building the AFST, Allegheny County provided space for legal experts, academics, community activists, and former foster youth to provide feedback and input into the system's construction and use. DHS held regular community meetings allowing legal experts, child welfare advocates, and community members to pose questions, raise concerns, and participate in the planning process of the county's new predictive analytics tool. Additionally, Allegheny County commissioned an independent ethical review and process and impact analyses to better understand and attempt to mitigate the harm associated with using an algorithm to aid child welfare decision making. DHS keeps the AFST as a county-owned system in order to maintain accountability to the public, and freely shares information about the methodology, algorithm, process, and evaluations. Public data, reports, and information on the use of the system are published online, allowing advocates, experts, academics, and the media to continuously comment, analyze, and scrutinize the workings of the system. Beginning in 2020, DHS officials are participating in a University of Pittsburgh-led task force to better understand how algorithms and predictive analytics are impacting county residents.

Community advocates have raised concerns over the AFST's potential for "poverty profiling"—over-surveillance of families who are more likely to come in contact with public services. Since the algorithm's inception, the county has seen a reduction in observed racial disparity among cases screened for further investigation. Importantly, families with low-risk cases are less frequently subjected to intrusive investigations, while high-risk cases are being investigated at higher rates. This statistic highlights the use of predictive analytics to help inform, rather than replace, human decision making.

While predictive analytics remains a common yet controversial practice, the development of the Allegheny Family Screening Tool has included current best practices for this work, including prioritizing community voices, transparent processes, and accepting critical feedback while being responsive to concerns. Most importantly, the risk assessment score is just one part of the screening process, with human decision making prioritized. The preliminary results provide optimism about the DHS's work toward a more equitable and more efficient child welfare system in Allegheny County.

### **Automating.NYC** by AISP with contributions from Deepra Yusuf, Elyse Voegeli, Akina Younge, & Jon Truong

**Who:** Automating.NYC

**Where:** New York City

**Organization Type:** University

**Domains:** All government agency-held data sources

**Goal/Impact:** To empower New Yorkers to advocate for automated decision systems that work to undo unjust systems instead of encoding inequality.

New York City convened the Automated Decision Systems Task Force to assess the use and proliferation of automated decision systems (ADS) across social and city services in May 2018. The Task Force, which was mandated under Local Law 49, required participants to develop a series of recommendations on the use of ADS and would include community outreach and input as part of this process. Recognizing that many community members were receiving competing messages about the use of ADS, four then-students at the Harvard Kennedy School designed and built [Automating.NYC](#)—a publicly accessible website—to help make conversations about ADS more accessible to community members prior to the first public engagement forum on April 30, 2019. It provides a framework to critique ADSs and breaks down machine learning jargon. The purpose is "to empower New Yorkers to advocate for ADSs that work to undo unjust systems instead of encoding inequality." Automating.NYC is not affiliated with the Task Force and was the subject of Aki Younge, Deepra Yusuf, Elyse Voegeli, and Jon Truong's master's thesis.

In January 2019, Younge, Yusuf, Voegeli, and Truong leveraged their design and data expertise to construct a website that was accessible, visually appealing, and engaging to support community members in asking informed questions about automated decision systems, and with the hope that future systems are built to benefit them. Members of the group were all previously involved in efforts to promote algorithmic justice and had been accumulating knowledge on NYC-specific ADS since fall of 2018. To build the website, they used information from investigative reporting, research, and open data as well as working with New York agencies to develop case studies. The use of clear formatting, nontechnical and accessible language, and a series of examples across different social services systems allows for a broad audience with varying experience or awareness of ADS.

To explain algorithms, the site applies the metaphor of purchasing avocados to that of a basic algorithmic decision tree, adapted from Cathy O'Neil's book *Weapons of Math Destruction*—a choice-making activity that is both widely applicable and topical. A "toy algorithm" built using data from the NYC Open Data Portal then offers users the opportunity "to help an imaginary fire department predict which buildings are at high risk for fire" using a series



of variables; it demonstrates how automated systems, though driven by data points, are always influenced by the human decision making around what goes into the algorithm. The site further breaks down how the design of an ADS can contribute to an unjust system based on: impact, bias, explainability, automation, and flexibility. Within each of these sections, the reader learns where to detect potential harm and how it relates to exacerbation and, often, automation of injustice.

Three local examples of the application of ADSs and related outcomes are then provided: NYC Public Library, Housing Authority, and Department of Education. The comprehensive brief and digestible summaries provide readers with context on how ADS impact their lives as both individuals and as part of their broader community. The final “Get Active” section offers concrete steps to build community, advocate to local government, and learn more about these issues. While the creators of Automating NYC have completed their thesis, the site remains operable and concepts remain applicable to algorithmic decision making as it functions in local governments across the United States.

## Racial Equity in Data Analysis: WORK IN ACTION

### #Change FocusNYC by Sarah Zeller Berkman

**Who:** #ChangeFocus NYC, Intergenerational Change Initiative, Administration for Children’s Services, & Youth Studies Programs at the CUNY School of Professional Studies

**Where:** New York City

**Organization Type:** Government Agency, University

**Domains:** Education, child abuse, preventive, detention

**Goal/Impact:** To better understand the experiences of NYC youth ages 14–21 who have navigated multiple city agencies, and develop policy or programmatic recommendations to benefit youth ages 14–21 who are cross-agency involved.

In 2016, the Administration for Children’s Services (ACS) in New York City received a grant to work with the National Council on Crime and Delinquency (NCCD) to engage in a data sharing initiative. ACS partnered with the NYC Department of Education (DOE) to examine indicators across their shared datasets. In an effort to ensure that an analysis of these indicators felt useful to both participating agencies and to community members that are impacted by these systems, the Educational Services Division of Youth & Family Justice formed a community engagement committee. The small group composed of agency leads, academics, and community activists was charged with creating opportunities for people outside ACS and DOE to weigh in and keep the agencies accountable.

The first gathering brought together about 20 people working in schools, community-based organizations, and health clinics to talk about the initiative and to articulate which variables might be of interest to look at across the two agency datasets. Participants also identified data that was of interest but that they did not think was captured in the indicators. This community gathering, in conjunction with meetings that included a broad group of stakeholders from city agencies, homed in on 40 indicators that NCCD would analyze related to education, child abuse, prevention, and detention.

Although the people engaged in these initial conversations had experience working with ACS, city agencies, or affected communities, the community engagement committee felt that those most impacted by the two participating agencies, namely young people, remained unrepresented. One of the co-chairs of the committee,

Sarah Zeller-Berkman, Director of Youth Studies Programs at the City University of New York (CUNY), who is a critical participatory action researcher, proposed that sustained engagement with a group of young people who had been in ACS care (preventive or juvenile justice) or young people who had fallen behind in middle school could generate data from other young people about the *lived* experience of being enmeshed in these systems. With the support of the community engagement committee and funding from one of its members, the #ChangeFocusNYC project<sup>54</sup> was developed in spring of 2017.

Over 140 people applied to be part of the first cohort of researchers working on the #ChangeFocusNYC project. The initial 15 youth were chosen for the Intergenerational Change Initiative (ICI) team to reflect a diversity of experiences, skills, and talents. Most youth team members had either fallen behind in school or had been involved with ACS through foster care or juvenile justice, while others were simply committed to making positive change using research. The young people partnered with Youth Studies master's students and the Director of Youth Studies Programs from the CUNY School of Professional Studies as well as a doctoral student from the CUNY Graduate Center to collaboratively design and implement this participatory action research study.

Teams of at least two youth facilitators and one adult conducted six focus groups in the Bronx, Manhattan, and Brooklyn with 68 participants between the ages of 14 and 24 to answer two main questions:

1. What are the experiences of NYC youth ages 14–21 who have dealt with multiple city agencies in their lives?
2. What are policy and/or programmatic recommendations that could benefit youth ages 14–21 who are dealing with multiple agencies in their lives?

Participants were recruited through research team contacts, community-based organizations, partner contacts, and community outreach, and represented a mix of youth on probation, youth who had fallen behind in school, and self-selected individuals who were at some point in direct contact with ACS.

The ICI team engaged in a deeply participatory process of iterative thematic coding that surfaced three categories, 10 themes, and 46 codes. The top six findings mentioned across all six focus groups were:

1. Agencies and programs serving the youth in our study insufficiently met their objectives.
2. Implementation of programs can at times be counterproductive to agency goals.
3. Interactions with agencies led youth in our study to feel unheard, stressed, and put down.
4. Youth would like to improve the professionalism of their service providers.
5. Youth desire personalized support and a sense of connection.
6. Youth want programming that meets their needs.

The group leveraged the expertise of research participation to develop recommendations that not only address the specific issues identified but also create processes that continually engage young people to affect the institutions that impact their lives. Included in the recommendations is youth involvement in shaping policies, developing monitoring mechanisms, training staff, and auditing outcomes. See the #ChangeFocusNYC research brief<sup>55</sup> for specific recommendations for ACS and DOE.

While the broader data sharing initiative was delayed for almost two years because of legal considerations, the ICI team distributed their findings in a variety of ways and with multiple audiences. The team created animated

54 See <http://www.intergenerationalchange.org/changefocusnyc.html>

55 See <http://www.intergenerationalchange.org/our-publications.html>



videos of the focus group data, a research brief, postcards displaying the findings, and interactive workshops using scenarios from the data. The community engagement committee reconvened in 2019 and is working with young people from the #ChangeFocusNYC project to share their findings and weigh in on the types of analysis ACS should conduct using the city data. Expanding the group that is making decisions about data use to include young people with lived experiences is an effective way to apply equity checks throughout the research and analysis process. While this project is ongoing, it illuminates how community-engaged initiatives help to ensure that those most impacted by systems are part of developing and using data to improve them.

## Racial Equity in Reporting and Dissemination: WORK IN ACTION

DataWorks NC by Libby McClure & John Killeen

**Who:** DataWorks NC

**Where:** Durham, North Carolina

**Organization Types:** Government agency, community-based organization

**Domains:** Housing, education, health, environment, infrastructure, community organizing

**Goal/Impact:** To serve neighbors and neighborhoods, nonprofits, and local governments by democratizing data to facilitate an empowered, productive, and equitable community.

DataWorks NC is a nonprofit, community-data intermediary in Durham, North Carolina. Their work is aimed at local neighborhoods, community organizers, and local government, with a priority for empowering people historically excluded from processes of power. One of DataWorks NC's main data projects is the [Neighborhood Compass](https://compass.durhamnc.gov/en),<sup>56</sup> a public-facing sub-county mapping tool of publicly available information, including administrative data. Additionally, they work with community groups, providing analysis and data support for their organizing efforts. Through their efforts, DataWorks NC has come to understand that while data can be a powerful tool for highlighting community concerns, it can also be used to undermine organizing efforts. Below, DataWorks NC documents some potential pitfalls in data reporting that re-create racism and antidotal, anti-racist reporting practices. Each anti-racist practice is accompanied by an example of work that demonstrates its value.

<sup>56</sup> See <https://compass.durhamnc.gov/en>

RACIST DATA REPORTING PRACTICE	ANTI-RACIST DATA REPORTING PRACTICE
<p>Using findings to justify harms or putting the burden on impacted communities to prove harm</p> <p>Reliance on causal inference and reactionary policy can force people to prove harm rather than requiring industries to prove safety</p>	<p>Promote solutions that change systems, not people</p> <p>Facilitate the development of policy by directly impacted communities</p>
<p>Focusing on deficits</p> <p>A lot of health research shows neighborhood disease rates without showing any neighborhood resources</p>	<p>Lead with assets and aspirations</p> <p>Convene collaborative, consensus-building sessions where people have space to speak what they care about most</p>
<p>Not responding to, being motivated by, or reporting back to communities</p> <p>Many health research topics that are driven by funding rather than community concerns</p>	<p>Prioritize relationships over numbers</p> <p>Take the time to understand community concerns, and only go where we're invited</p>

Motivated by community conversations regarding the state of evictions in Durham, DataWorks NC convened a series of events centered on changes in ownership of rental properties and eviction trends. At *Who Owns Durham?* events, DataWorks NC displayed relevant data in a “gallery walk.” Some pieces were participatory, like a map of evictions in Durham County on which attendees could mark where they live. Conversations evolved into strategy sessions where tenants were given a platform to collaboratively generate ideas to prevent displacement. The work culminated in a [strategy document](#)<sup>57</sup> that considers ways to reduce harm among individuals who have been evicted, the root causes of displacement, and the potential avenues for and barriers to making change. DataWorks NC also wrote a “[Quick Takes](#)”<sup>58</sup> document—a simple, two-sided piece of paper with distilled, clarified statements documenting trends and disparities in wages, rent, and corporate ownership. They have distributed it among organizers in Durham who are using it in their work right now.

Participants also frequently referenced an imbalance in awareness of market trends, practices, and laws between tenants and corporate owners, so DataWorks NC created a [webpage](#)<sup>59</sup> for tenants and organizers to begin to address this inequity. The site documents experiences and disparities related to evictions, offers resources for tenants facing potential eviction, and includes easily used data support for tenant advocacy.

DataWorks NC also hosts regular “research pod” meetings, where data collection becomes organizing. Groups of residents meet at the DataWorks office to investigate corporate ownership stakeholders and patterns in Durham. This work is ongoing.

<sup>57</sup> See [http://bit.ly/anti\\_displacement](http://bit.ly/anti_displacement)

<sup>58</sup> See [https://docs.google.com/document/d/1TQDz96jww-wisizUinHN1u8JKGBLrRV3hbqCXwvz\\_A4/edit](https://docs.google.com/document/d/1TQDz96jww-wisizUinHN1u8JKGBLrRV3hbqCXwvz_A4/edit)

<sup>59</sup> See <https://dataworks-nc.org/evictions/>

## Metriarch Data Lookbook and Lady Charts Event by Emma Swepston, Laura Bellis, & Brandy Hammons

**Who:** The Take Control Initiative

**Where:** Tulsa, Oklahoma

**Organization Types:** Government agency, community-based organization

**Domains:** Health care access, Supplemental Nutritional Assistance Program (SNAP), economic factors, adolescent health, education, maternal health, mental health, incarceration, intimate partner violence

**Goal/Impact:** To normalize and broaden the conversations around women's health in Oklahoma through data storytelling, resource curation, and interactive outreach events.

In 2018, the [Take Control Initiative](https://www.takecontrolinitiative.org/)<sup>60</sup> began searching for partners for a new statewide data collaborative and convened Metriarch, a data haven for Oklahoma women. Metriarch aims to normalize and broaden the women's health conversation in Oklahoma through data storytelling, resource curation, and interactive outreach events. The data collaborative is a partnership between 30 policy, direct service, community coalition, and nonprofit organizations in the state. Using a group-centered leadership approach, Metriarch has been able to center historically marginalized voices and create conversations around pressing issues related to women's health.

Before Metriarch, Oklahoma organizations felt restricted by the limited publicly available data (typically produced annually). Agencies and individuals across the community also reported feeling confused and overwhelmed by the ways data were being reported. Metriarch approached these challenges head-on and took an innovative approach to reporting, access, and outreach that makes data usable by all Oklahoma residents and service providers.

Metriarch sought to expand the use of data as a tool to promote community change. From this vision, Metriarch released the [Lookbook](#)<sup>61</sup>—a straightforward dissemination of data that uses a narrative framework to contextualize data in a way that makes it relevant and useful for community members, service organizations, academics, and policy makers. Additionally, Metriarch connects with the community through innovative approaches to outreach such as their [Lady Charts](#)<sup>62</sup> program. In fall of 2019, Lady Charts—a women's health data jam—brought together data mavens, advocates, and community leaders to celebrate, educate, and engage in conversation around women's health. Lady Charts has brought new partners to the table and furthered Metriarch's work by facilitating meaningful dialogue around what data are currently measured, as well as what stories are being left out of the data. Metriarch has taken steps to ensure that Oklahomans have access to their data through a series of online web tools, dashboards, references, and a legislative tool. The [legislative tool](#)<sup>63</sup> is a visualization tool tracking bills through the Oklahoma State House and Senate. The tool allows collaborative members to come together, look at bills, and highlight pertinent policy issues and leverage data for advocacy work.

Without proper access, analysis, and dissemination, community stories can be lost in the numbers. Metriarch's innovative approach is changing the way data can be shared, analyzed, and used to facilitate meaningful conversation and societal change. Their work has empowered the community and helped ensure that the data collected by organizations further the service missions of community collaborators.

60 See <https://www.takecontrolinitiative.org/>

61 See <https://www.metriarchok.org/catalogs/>

62 See <https://www.eventbrite.com/e/lady-charts-2019-tickets-71130803147#>

63 See <https://www.metriarchok.org/legislative-tracker/>

## City of Asheville, Mapping Racial Equity by Christen McNamara & Kimberlee Archie

**Who:** Office of Equity & Inclusion, GIS

**Where:** Asheville, NC

**Organization Types:** Government Agency

**Domains:** Equity and inclusion, training, GIS, spatial data analysis, open data

**Goal/Impact:** To advance racial equity and create more inclusive prosperity, where the well-being for all improves, by including equity data in GIS to enable the City of Asheville to make better decisions about planning, affordable housing, and economic and educational growth.

In January 2016, the Asheville City Council created a 20-year vision for the city that included the statement, “cultural diversity and social and economic equity are evident in all that we do.” The Office of Equity & Inclusion recognizes the different places, opportunities, obstacles, and resource needs that impact people’s ability to live and thrive. In recent years, Asheville has experienced growth in population, tourism, and economic activity. As an unintended consequence of this growth, low-income populations and BIPOC have faced new barriers to opportunity, such as gentrification and displacement.

The City of Asheville (COA) is focusing on eliminating racial inequities by developing policies and practices that benefit people who have historically been held back and neglected in response to these issues. COA believes that creating these policies will help everyone thrive. The “why” for this work stems from a value of equity and inclusion as an organization. The Office of Equity & Inclusion recognizes the local government’s role in creating and maintaining the disparities experienced by BIPOC and people living in poverty. They also recognize the city’s responsibility to change how it operates in order to redistribute and share power to change community conditions.

To support this work, the Office of Equity & Inclusion and the city’s GIS team partnered to create a story map: [Mapping Racial Equity in Asheville, NC](#).<sup>64</sup> This project uses agency-held administrative data to visualize historic, racist policies and the racialized impacts the policies have in Asheville geographically. The Office of Equity & Inclusion has presented their story map to local community groups and teams and solicited feedback to inform future agency projects.

Linking equity data to location through GIS software has allowed the Office of Equity and Inclusion to better understand the connection between racialized policies and physical location. Historically, policies like Home Owners Loan Corporation (HOLC) red-lining and urban renewal have used location-based restrictions to limit access to opportunities, such as home ownership, for BIPOC. Mapping equity data has allowed the Office of Equity and Inclusion to study the generational cycles of racially based economic and social disadvantages created by location-based policies and help correct their racialized consequences.

Further, equity data is held in the city’s GIS, which makes it accessible to individuals and organizations at all levels of government and community work. Community members and groups working to advance racial equity can access these data through the Storymap or City Data Resource to inform their work. Likewise, every city department can use these data to make decisions about “who” and “where” to implement programming. For instance, the Parks and Recreation Department can see the neighborhoods with large, racialized gaps in educational achievement and provide supportive educational programming, such as tutoring, in their afterschool programs and camps in these areas. The Planning Department can use the Neighborhood Change Index to apply anti-gentrification and anti-displacement strategies. Public Safety continues to use the Open Data Portal to build trust with communities where trust is currently lacking. Affordable Housing staff can use urban renewal data to make decisions about the “where” of their initiatives and when the “who” requires additional access and opportunities.

64 See <https://avl.maps.arcgis.com/apps/Cascade/index.html?appid=10d222eb75854cba994b9a0083a40740/>

Additionally, these data provide Asheville with important insights on neighborhood needs when planning future capital projects and investments. Using the framework of the Government Alliance on Race and Equity (GARE) Racial Equity Toolkit, the COA Equity Core Team developed a Budget Equity Tool for budget enhancements. When the city's budget is being reviewed, equity data provide insight into "where" and "who" needs funding in order to eliminate race-based inequities. Overall, equity data provide metrics to help measure the effectiveness of new policies, programs, and investments aimed at advancing racial equity in COA.

The Office of Equity & Inclusion considers Mapping Racial Equity to be a living project. As new data become available and new organizational needs arise, the tool will continue to be updated and improved. The Mapping Racial Equity in Asheville, NC project won the Herb Stout Award for Visionary Use of GIS at the North Carolina GIS Conference in March 2019 and continues to receive recognition throughout North Carolina and across the United States.

## Appendix II: Terms

**Administrative Data:** Data collected during the routine process of administering programs.

**Administrative Data Reuse:** Using administrative data in a way not originally intended, e.g., for research.

**Bias:** Interference in outcomes by predetermined ideas, prejudice or influence in a certain direction. Bias can be both implicit and explicit. Implicit bias refers to the attitudes or stereotypes that affect our understanding, actions, and decisions in an unconscious manner. All data is inherently biased due to sampling, collection, interpretation, and dissemination.  
See <http://kirwaninstitute.osu.edu/research/understanding-implicit-bias/>

**Community:** A place where people reside and interact, or a larger system that people are a part of (e.g., youth in foster care). We consider the members of a community to have similar characteristics, experiences, or interests.  
See <https://www.childtrends.org/publications/a-guide-to-incorporating-a-racial-and-ethnic-equity-perspective-throughout-the-research-process>

**Community Engagement:** The process of working collaboratively with and through groups of people to address issues affecting the group's well-being. Community engagement should include authentic processes at all stages of a project—particularly the beginning—that partner agencies directly with community stakeholders. Centering the community in agency work is necessary to achieve long-term and sustainable outcomes.

**Consent:** Explicit permission regarding the collection, storage, management, and use of personal information. Individuals can give active (i.e., opt-in) or passive consent (i.e., implicit or opt-out). Consent must be freely given, specific, informed, and unambiguous.  
See [https://fpf.org/wp-content/uploads/2018/09/FPF-AISP\\_Nothing-to-Hide.pdf](https://fpf.org/wp-content/uploads/2018/09/FPF-AISP_Nothing-to-Hide.pdf), pg 21

**Data:** Information collected to help decision making

**Data Ethics:** Ethical issues that arise from recognizing that the evaluation of data collection, sharing, and use practices has the potential for adverse impacts. In order to minimize adverse impacts of data processes, ethical concerns should be addressed at all stages of the data life cycle:

- Data stewardship: the collection, maintenance, and sharing of data
- Information generation: using data to create products and services, analysis and insights, and stories or visualizations
- Data decision making: bringing together data with practical experience and knowledge to make informed policy or practice decisions

See Open Data Institute, <https://theodi.org/article/data-ethics-canvas/>

**Data Governance:** The policies and procedures that determine how data are managed, used, and protected.

**Data Infrastructure:** Approach for using, storing, securing, and interpreting data. Includes physical elements, such as storage approaches, and intangible elements such as software.

**Data Privacy:** A fundamental right, guaranteed by law, that provides individuals control over the collection, use, and dissemination of their data. Agencies must put procedures in place that protect how personal information flows, protect the freedom of thought and exploration; and protect one's dignity and reputation.

See [https://fpf.org/wp-content/uploads/2018/09/FPF-AISP\\_Nothing-to-Hide.pdf](https://fpf.org/wp-content/uploads/2018/09/FPF-AISP_Nothing-to-Hide.pdf), Appendix A

**Equity:** The elimination of privilege, oppression, disparities, and disadvantage to co-create a just, fair, and inclusive society in which all can participate, prosper, and reach their full potential.

See <https://www.policylink.org/about-us/equity-manifesto>

**Racial Equity:** The condition where one's racial identity no longer influences how one fares in society. This includes the creation of racially just policies, practices, attitudes, and cultural messages, and the elimination of structures that reinforce differential experiences and outcomes by race.

See <http://racialequitytools.org/glossary#racial-equity>

## Racism

**Individual Racism:** The beliefs, attitudes, and actions of individuals that support or perpetuate racism in conscious and unconscious ways.

See <http://racialequitytools.org/fundamentals/core-concepts/racism>

**Institutional Racism:** The ways in which policies, procedures, and practices of parts of systems (e.g., schools, courts, transportation authorities) or organizations influence different outcomes for different racial groups.

See <http://racialequitytools.org/fundamentals/core-concepts/racism>

**Structural Racism:** The normalization and legitimization of an array of dynamics—historical, cultural, institutional and interpersonal—that routinely advantage Whites while producing cumulative and chronic adverse outcomes for people of color.

See <http://racialequitytools.org/glossary#structural-racism>

**Social Justice:** The proactive reinforcement of policies, practices, attitudes, and actions that produce equitable power, access, opportunities, treatment, impacts, and outcomes for all.

See <http://racialequitytools.org/glossary#racial-equity>

## Appendix III: Expanded Civic Data User Profiles

### Members of institutions

- **University-based researcher** who works with government and nonprofit agencies to support research and evaluation efforts that lead to improved policy and outcomes.
- **State government agency situated administrator** who works across state agencies to better drive integrated service delivery and outcome improvement through data and policy.
- **State government/local government agency of technology administrator** who works across agencies to support data infrastructure and use.
- **City/county government agency situated administrator** that works across agencies to better drive integrated service delivery and outcome improvement through data and policy.
- **Foundation staff** who work to better understand and support investments that lead to quality research and evaluation, and ultimately improved outcomes for families and communities.

#### Questions to guide members of institutions

- Historically, how has our use of data impacted disenfranchised communities we seek to serve?
- What must we do differently to center equity?

### Community representatives

- **Community advocate** who works to build transparency and center equity within government-led projects.
- **Community leader/members** who live within, receives, is directly impacted by services and data collection, including members of over-surveilled communities.
- **Local backbone organization** (community-based nonprofit, NGO, Children Services Council) **administrator** who works across agencies to better drive integrated service delivery and outcome improvement through data and policy.

#### Questions to guide community representatives

- What do we not have access to that would have made my work more impactful?
- How has data been used against me and my work?



## Bridges between community and institutions

- **Service provider/social worker/case worker** (any agency/entity receiving direct funding from government to provide services) who works across agencies to drive integrated service and improved outcomes (homelessness service providers, social workers, etc.).
- **Applied researcher** who works with government and nonprofit agencies.
- **Consultant** who works with government and nonprofit partners to improve social outcomes.

### Questions to guide bridges between community and institutions

- How have we been successful in using data to move individual- and population-level outcomes?
- What are the opportunities to address equity?
- What are the assumptions we are operating under, particularly as related to data, class, and race?

## Appendix IV: Resources

### General Resources

Annie E. Casey Foundation. (2014). [Race Equity and Inclusion Action Guide](https://www.aecf.org/resources/race-equity-and-inclusion-action-guide/). <https://www.aecf.org/resources/race-equity-and-inclusion-action-guide/>

Annie E. Casey Foundation. (2018). [Race Equity Crosswalk Tool](https://www.aecf.org/m/blogdoc/aecf-raceequitycrosswalk-2018.pdf), Annie E. Casey Foundation. <https://www.aecf.org/m/blogdoc/aecf-raceequitycrosswalk-2018.pdf>

Chicago Beyond. (2019). [Why am I Always Being Researched?](https://chicagobeyond.org/wp-content/uploads/2019/05/ChicagoBeyond_2019Guidebook.pdf) [https://chicagobeyond.org/wp-content/uploads/2019/05/ChicagoBeyond\\_2019Guidebook.pdf](https://chicagobeyond.org/wp-content/uploads/2019/05/ChicagoBeyond_2019Guidebook.pdf)

Coalition of Communities of Color. (2014). [Tool for Organizational Self-Assessment Related to Racial Equity](https://www.coalitioncommunitiescolor.org/research-and-publications/ccorgassessment). <https://www.coalitioncommunitiescolor.org/research-and-publications/ccorgassessment>

Equity in the Center, ProInspire. (2019). [Awake to Woke to Work: Building a Race Equity Culture](https://www.equityinthecenter.org/wp-content/uploads/2019/04/Equity-in-Center-Awake-Woke-Work-2019-final-1.pdf). <https://www.equityinthecenter.org/wp-content/uploads/2019/04/Equity-in-Center-Awake-Woke-Work-2019-final-1.pdf>

Nelson, J., Spokane, L., Ross, L., & Deng, N. (2015). [Advancing Racial Equity and Transforming Government: A Resource Guide to Put Ideas into Action](https://racialequityalliance.org/wp-content/uploads/2015/02/GARE-Resource-Guide.pdf). Local and Regional Government Alliance on Race & Equity. <https://racialequityalliance.org/wp-content/uploads/2015/02/GARE-Resource-Guide.pdf>

Nelson, J., & Brooks, L. (2015). [Racial Equity Toolkit: An opportunity to operationalize equity](https://www.racialequityalliance.org/wp-content/uploads/2015/10/GARE-Racial-Equity-Toolkit.pdf). Local and Regional Government Alliance on Race & Equity. <https://www.racialequityalliance.org/wp-content/uploads/2015/10/GARE-Racial-Equity-Toolkit.pdf>

Lewis, T., Gangadharan, S. P., Saba, M., Petty, T. (2018). [Digital defense playbook: Community power tools for reclaiming data](https://www.odbpject.org/wp-content/uploads/2019/03/ODB-DDP_HighRes_Spreads.pdf). Detroit: Our Data Bodies. [https://www.odbpject.org/wp-content/uploads/2019/03/ODB-DDP\\_HighRes\\_Spreads.pdf](https://www.odbpject.org/wp-content/uploads/2019/03/ODB-DDP_HighRes_Spreads.pdf)

United Kingdom Department for Digital, Culture, Media & Sport. (2018). [Data Ethics Workbook](https://www.gov.uk/government/publications/data-ethics-workbook/data-ethics-workbook). <https://www.gov.uk/government/publications/data-ethics-workbook/data-ethics-workbook>

### Specific to Algorithms

Anderson, D., Bonaguru, J., McKinney, M., Nicklin, A., Wiseman, J. (2018). [Ethics & Algorithms Toolkit](https://ethicstoolkit.ai/). Center for Government Excellence, Johns Hopkins University. <https://ethicstoolkit.ai/>

Diakopoulos, N., Friedler, S., Arenas, M., Barocas, S., Hay, M., Howe, B., Jagadish, H. V., Unsworth, K., Sahuguet, A., Venkatasubramanian, S., Wilson, C., Yu, C., & Zevenbergen, B. (2017). [Principles for accountable algorithms and a social impact statement for algorithms](https://www.fatml.org/resources/principles-for-accountable-algorithms). FAT/ML. <https://www.fatml.org/resources/principles-for-accountable-algorithms>

Katell, M., Young, M., Dailey, D., Herman, B., Guetler, V., Tam, A., Binz, C., Raz, D., & Krafft, P.M. (2020). [Toward Situated Interventions for Algorithmic Equity: Lessons from the Field](#), Association for Computing Machinery. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. <https://dl.acm.org/doi/abs/10.1145/3351095.3372874>

Kingsley, C. (2017). [First, Do No Harm: Ethical Guidelines for Applying Predictive Tools Within Human Services](#). MetroLab Network. [https://metrolabnetwork.org/wp-content/uploads/2017/09/Ethical-Guidelines-for-Applying-Predictive-Tools-within-Human-Services\\_Sept-2017.pdf](https://metrolabnetwork.org/wp-content/uploads/2017/09/Ethical-Guidelines-for-Applying-Predictive-Tools-within-Human-Services_Sept-2017.pdf)

Richardson, R. ed. (2019). [Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force](#). AI Now Institute. <https://ainowinstitute.org/ads-shadowreport-2019.pdf>

## Specific to Foundation Staff

Lief, L. (2020). [How Philanthropy Can Help Lead on Data Justice](#), Stanford Social Innovation Review. [https://ssir.org/articles/entry/how\\_philanthropy\\_can\\_help\\_lead\\_on\\_data\\_justice#](https://ssir.org/articles/entry/how_philanthropy_can_help_lead_on_data_justice#)

## Specific to Researchers

Andrews, K., Parekh, J., Peckoo, S. (2019). [A guide to incorporating a racial and ethnic equity perspective throughout the research process](#). Child Trends. [https://www.childtrends.org/wp-content/uploads/2019/09/RacialEthnicEquityPerspective\\_ChildTrends\\_October2019.pdf](https://www.childtrends.org/wp-content/uploads/2019/09/RacialEthnicEquityPerspective_ChildTrends_October2019.pdf)

Kirkland, D. (2019). [No Small Matters: Reimagining the Use of Research Evidence From A Racial Justice Perspective](#). William T. Grant Foundation. <http://wtgrantfoundation.org/digest/no-small-matters-reimagining-the-use-of-research-evidence-from-a-racial-justice-perspective>

Actionable Intelligence for Social Policy

University of Pennsylvania

3701 Locust Walk, Philadelphia, PA 19104

215.573.5827 [www.aisp.upenn.edu](http://www.aisp.upenn.edu)

