



Envisioning a
New Future

Building Trust for Data Use

By Leah Hendey and
Kathryn L.S. Pettit,
Urban Institute and
the National Neighborhood
Indicators Partnership

**FOR THE DATA FUNDERS
COLLABORATIVE**

Acknowledgments

This paper was produced with and for the [Data Funders Collaborative](#), a community of philanthropic practitioners working together to support and lead dialogue and action around the collection and ethical use of data to empower communities, increase equity, and improve outcomes in the education and social services sectors. We want to especially acknowledge Aimee Rogstad Guidera (Data Funders Collaborative Coordinator) and Chris Kingsley (Annie E. Casey Foundation) for their substantive input and guidance and for sharing the many examples from the Data Funders Collaborative. We thank the Data Funders Collaborative leadership group for this series of papers for their thoughtful comments and encouraging words: George Hobor (Robert Wood Johnson Foundation), Michelle Shevin and Maia Woluchem (Ford Foundation), and Jon Sotsky (Overdeck Family Foundation). We also benefitted from the feedback shared by other members of the collaborative.

The views expressed are those of the author/authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples.

About the Data Funders Collaborative

DFC: The Data Funders Collaborative (DFC) is a partnership of leading philanthropic organizations working together to support learning, discovery and action focused on the ethical collection, protection and use of data. DFC members recognize the power of information to both inform better decision-making and confront entrenched inequality. Our aim is to help communities become better informed and able to achieve equitable outcomes in education, health and other social services sectors. For more information on the DFC, please contact Aimee Rogstad Guidera, Data Funders Collaborative Coordinator at Aimee@GuideraStrategy.com or 703-303-6912

About the Urban Institute

The nonprofit Urban Institute is a leading research organization dedicated to developing evidence-based insights that improve people's lives and strengthen communities. For 50 years, Urban has been the trusted source for rigorous analysis of complex social and economic issues; strategic advice to policymakers, philanthropists, and practitioners; and new, promising ideas that expand opportunities for all. Our work inspires effective decisions that advance fairness and enhance the well-being of people and places. Since 1996, Urban Institute has coordinated the [National Neighborhood Indicators Partnership](#) (NNIP). NNIP is a learning network connecting partner organizations in more than 30 cities that ensure their communities have access to data and the skills to use information to advance equity and well-being across neighborhoods.

Contents

Foreword	1
Introduction	2
Methods and Definitions	4
Failures in Data Use and Integration	5
Rights and Harms	8
Approaches to Building Trust	10
Enact and Refine Laws and Regulations Related to Data	11
Apply Technical Solutions to Expand and Control Data Access	12
Increase Community Data Capacity	13
Establish and Enhance Governance for Data and Data Systems	15
Roles for Philanthropy	17
Ways to Advance the Field	17
Considerations for Investments	18
Areas for Investment	19
Conclusion	20
Appendix A	21
References	30

Foreword

This paper is the culminating public-facing product of a three-part landscape review around integrated data in education and social services. The authors produced this with and for the [Data Funders Collaborative](#) (DFC), a community of philanthropic practitioners working together to support and lead dialogue and action around the collection and ethical use of data to empower communities, increase equity, and improve outcomes in the education and social services sectors.

Given the country's and the philanthropic community's increased focus on racial equity and social justice, the DFC is eager to better understand how data can be used to improve outcomes for all people, especially those traditionally underserved or harmed by its misuse.

When the DFC began convening four years ago, there was a particular interest in data linked and shared across disparate datasets, the data necessary to solve the knottiest challenges in the social services sector. The DFC is heartened by the positive use of these integrated data as a tool of empowerment and improvement but is also aware of the real and potential harm of these data to the people they are meant to serve. Now is an opportune time to understand the current state of integrated data in the social services sector, study the promise and perils of its use, identify timely opportunities to strengthen the data ecosystem, and to offer actions for the funder community to consider. One of the most important insights highlighted in all three papers is the need to better engage the communities this work aims to serve; without this crucial change of focus, there is little chance of harnessing the power of data to contribute to real, sustainable change and improvement in these communities.

The first two papers of the Integrated Data Landscape Review were created as working documents to inform the ongoing discussions, learning and strategies of the Data Funders Collaborative. Short descriptions of these two internal papers follow:

- ***Integrated Data Systems: A Review of the Existing Landscape*** provides a snapshot of current progress building and supporting integrated data systems; distills the components, capacity, and conditions required to nurture the growth and effectiveness of these systems; and offers considerations for funder action. (*Contributing author: Bi Vuong at Project Evident.*)
- ***Strengthening Integrated Data Capacity: The Federal Landscape*** captures past efforts and current opportunities to change how federal, state, and local governments use integrated data to improve the effectiveness and efficiency of a broad range of government functions. (*Author: Kathy Stack.*)

This paper, ***Envisioning a New Future: Building Trust for Data Use***, written by Kathy Pettit and Leah Hendey at the Urban Institute's National Neighborhood Indicators Partnership, while developed with and for the DFC, has a broader potential audience. It aims to provide a common language for funders regarding the opportunities and challenges to improving engagement and building trust with communities and offers a range of approaches to consider for future investments. The DFC looks forward to sharing this paper with the broader philanthropic community, expanding conversations, and learning together.

For more information on this landscape review or the DFC, please contact DFC Coordinator, Aimee Rogstad Guidera, at Aimee@GuideraStrategy.com.

Introduction

More state and local government agencies and nonprofits are publishing [open data](#), using data in their everyday operations, and integrating data across systems to better serve children, young people, adults, and families. The number of integrated data systems has grown, and they are increasingly used by the public and social sector in policymaking, social impact finance projects, collective impact initiatives, and responses to the COVID-19 pandemic. We believe data are a powerful tool for people to use to improve lives, strengthen families, and build communities of opportunity (box 1).

“ We believe data are a powerful tool for people to use to improve lives, strengthen families, and build communities of opportunity. ”

However, in creating integrated data systems and embedding data use in agencies, system designers, funders, and users paid insufficient attention to building trust with communities, particularly those

represented in state and local administrative data. This dynamic is no longer tolerable or productive. Persistently harmful data practices by both public and private sectors continue to erode trust, especially among Black, Indigenous, and other people of color, including Asian, Hispanic, and Pacific Islander people.¹ People represented in data are rarely included in the governance of those data and data systems or consulted on the interests and needs those data could be used to support. This lack of engagement and trust compromises investments in data systems and progress in stakeholders' use of data, as communities rightfully oppose data efforts that are not transparent and inclusive, as the case of the nonprofit organization, inBloom, demonstrates.

In 2013, with the support of the Gates and Carnegie Foundations, inBloom launched its trademark centralized education platform for curricula, data sharing, and learning apps. The \$100 million initiative aimed to improve interoperability, security, reporting, and data sharing to inform instruction and improve outcomes. However, poor communication, a lack of transparency, and a shift in public perceptions about data use in education, coinciding with a series of

BOX 1

How Data Can Benefit Communities

Funders, practitioners, and policymakers can use data and integrated data systems to achieve the following:

- improve operational efficiency and service delivery through better program management and resource allocation
- design, assess, and improve policy, including by capturing costs and benefits across service systems
- design, assess, and improve programmatic investments, analyzing data for overall outcomes

Researchers and people and communities outside of government can use individual-level data or summarized program data to facilitate the following:

- make informed decisions for themselves and their families about schools, health care, and other services
- act collectively to advocate for improved policies and programs, assess government investments, and set priorities for community action

¹ We use the terms Black, Indigenous, and people of color to acknowledge the systemic racism that continues to affect Black and Indigenous people in ways that differ from other people of color, including in their experiences of data being weaponized against them. We acknowledge this language may not reflect how people describe themselves.

private-sector data breaches, led to backlash over inBloom's intended integration and use of student data. These events launched a student data privacy movement and led to the dissolution of inBloom a year later (Bulger et al. 2017). A lack of engagement and trust building is also why the field has not seen the potentially transformative outcomes of effective data use.

In this paper, we aim to provide a common language for funders to surface assumptions about rights and harms related to data use and to propose an overarching framework of ways to build trust in how

data can guide future discussions and investments. We first describe major critiques of current data and data integration efforts and individual and collective rights to be protected and advanced. With this foundation, we offer approaches to building trust with communities and relevant examples. We conclude with a discussion on areas ripe for time and funding investments. Our primary audience is the Data Funders Collaborative and their members, though the concepts presented are relevant to philanthropy generally and to those invested in promoting data use to improve lives and community outcomes.

Methods and Definitions

For this paper, we drew from the valuable perspectives and publications of the many organizations advancing thinking on using data for the public good and centering equity and people in data practice, including guidance on when data should not be used. We also incorporated examples and viewpoints from conversations with the Data Funders Collaborative. Finally, we leveraged our insights from more than two decades of our coordinating the National Neighborhood Indicators Partnership, a learning network of independent organizations that ensure their communities have access to data and the skills to use information to advance equity and well-being across neighborhoods. The approaches to building trust presented later are our syntheses of these diverse sources and provide a framework to help foundations consider building trust with communities for data use when they assess their current investments and think about future ones.

Since *using data* means different things to different people, it is important to clarify scope and definitions for this paper. First, we include any data use and sharing by the government or social sector, whether part of an integrated data system or not. The public does not easily differentiate between the use of a single dataset and linked data.² And, more importantly, stewards and other stakeholders need to build people's trust and address concerns about the collection and use of data generally, as well as about the additional risks and benefits of linked data. For shorthand, this paper will refer to "data use and integration," but the phrase is meant to encapsulate the processes involved at each stage of the data cycle, from acquisition and sharing to processing and analysis, dissemination, and disposition.

This paper focuses on data use and sharing by government agencies or social sector organizations—

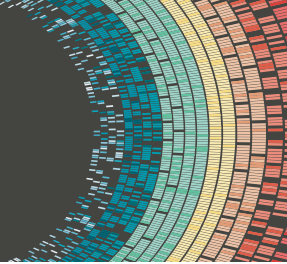
which primarily relates to state and local administrative data but may also include proprietary data and systems. We know the private sector also collects data on people, and firms are not often transparent about how they use and share those data. There are important conversations happening on the risks and benefits of the private sector's collection and use of data and how people can gain more control of their own data³ (see for example: Moss and Metcalf 2020, Sparapani and Sherman 2021, Healthcare Information and Management Systems Society 2021, and Mozilla Insights, van Geuns, and Brandusescu 2019). Though we do not address private-sector regulation in this paper, the sector's behavior affects public opinion and the environment for all data discussions, and data users should understand the latest issues and trends.

Finally, when we discuss *communities* in this paper, we are generally referring to all people living in a particular place, especially those represented in data. Widespread public trust for data use is needed from people of all income levels and races. However, we want to emphasize that Black, Indigenous, and people of color have been the victims of harmful uses of data and historically excluded in the governance of data and data systems, both of which continue today. They are also overrepresented in administrative data and continue to bear disproportionate risk of harm, as discrimination and structural racism baked into social service and other systems has limited their opportunities and resulted in differential treatment by race. Given this context, all stakeholders should prioritize building trust with Black, Indigenous, and people of color and center their experiences as we seek to improve how the field uses data for public good.

2 Topos Partnership, "Promoting Data Sharing Approaches" (Data Funders Collaborative: 2019), unpublished internal document.

3 HIMSS (Healthcare Information and Management Systems Society), ["Uncovering and Removing Data Bias in Healthcare,"](#) HIMSS: April 6, 2021.

Failures in Data Use and Integration



Box 1 on page 2 summarizes ways people can use data to improve individual outcomes and benefit communities. There are many examples that have demonstrated these positive uses of data. However, advocates and other stakeholders have voiced valid criticisms about the use and sharing of data by the government and social sector. These concerns stem in large part from top-down development of data collection and governance and a lack of engagement of people represented in the data and those affected by its use. They are perceived failures of laws and governance structures as well as of how data are used by local and state agencies. Over time, these concerns have eroded people's trust that data will be used responsibly and for the public good.

These concerns include:

1. Agencies do not have the tools to sufficiently protect privacy with the current legal, regulatory, and governance frameworks.
2. Communities that have been harmed or marginalized are not given authentic opportunities to protect themselves from unwanted data use or to be involved in decisions about uses.
3. People who could use data to make personal decisions, improve programs and services, or advocate for effective and equitable government do not have data in the ways they need.

We outline these concerns below and provide illustrative examples and organizations leading the discourse.

Concern #1: Agencies do not have the tools to sufficiently protect privacy with the current legal, regulatory, and governance frameworks.

Government agencies have developed these frameworks for the purposes of agency operations or reporting requirements, typically without input from anyone outside government. This lack of engagement, combined with high-profile data breaches and unethical data sharing (box 2), has resulted in public distrust in privacy protections. Though much of the mainstream discussion around data privacy relates to consumer rights in private-sector data use and technology firms' behavior, the concerns spill over onto government administrative data and use. People are seeking increased transparency from agencies about why data are collected and how data are used and shared, and they want to be involved in

developing the frameworks to ensure their privacy is prioritized.

In addition, in the US, government regulatory and governance frameworks have not kept up with the expanding types of data available. For example, unlike

BOX 2

Inappropriate Data Sharing Erodes Trust

In Pasco County, Florida, the sheriff's office obtained student education records without parental consent and no valid exception to Family Educational Rights and Privacy Act (FERPA) to generate a list of students who might become "potential criminals" and are "destined to a life of crime." The data sharing was supposedly intended to support school resource officers' interventions but was used for law enforcement, including for off-campus interventions.

“Once your information is out there, it doesn't matter if you want to share it or don't want to share it, somebody's going to have access to it.”

Topos: [Promoting Data Sharing Approaches](#)

with traditional administrative data that cover people who receive a service from the government and are likely aware their information is being retained and used in some way, some agencies use emerging private-sector data sources, like facial recognition software, to capture people who have not provided consent and who are not receiving any benefits. For

example, Clearview, a facial recognition software that scraped online photos to sell to law enforcement agencies, is now [banned in Canada](#). Both the [Center for Democracy & Technology](#) and the [Future of Privacy Forum](#) provide thought leadership on data privacy and focus on the rights of individuals.

Concern #2: Communities that have been harmed or marginalized are not given authentic opportunities to protect themselves from unwanted data use or to be involved in decisions about uses.

Another failure of existing data practice frameworks is the lack of avenues for people and communities who have been harmed by data or marginalized by government to authentically participate in data governance. This failure disproportionately affects Black, Indigenous, and people of color and their communities. Organizations providing leadership in this area believe privacy protections need to be improved but take a view that beyond an individual right to privacy, there are civil rights for groups of people at stake that are needed to rectify and prevent harms. The case study of Ramsey County, Minnesota, highlights this idea: government agencies (and funders), even if well intentioned, cannot see the risks and benefits of data sharing in the same way as people whose lives will be affected. Ramsey County, the City of St. Paul, and St. Paul City Schools drafted a joint powers agreement to enable cross-agency data

sharing and to improve service communication and coordination to reduce negative outcomes for young people. Though Ramsey County had convened a series of community engagement sessions around school violence, they had never discussed an integrated data system or advanced analytics as solutions with community members. Local organizations united as the [Stop the Cradle to Prison Algorithm Coalition](#) to prevent the data sharing. The coalition noted that the data reflected structural racism and bias against students of color, there was no mechanism for community oversight, and the community would not trust services to be provided based on the data. The coalition surfaced valid concerns and brought to light that supporters of the agreement hadn't considered all the potential negative consequences or articulated how they would protect the data.⁴

“ Our elected officials had no idea what they agreed to, and the implications that could come from it. Data is not bad, but data without any kind of oversight that includes the community does not benefit us. ”

— Marika Pfefferkorn, [Twin Cities Innovation Alliance](#)

Source: [TwinCities.com](#)

“ The political process that created the data sharing agreement excluded communities and families and could have resulted in serious harm. This process should not have happened, and communities need to know it won't be repeated. ”

— Laura LeBlanc, InEquality

Source: [Dignity in Schools](#)

⁴ “[Defeating the JPA](#)” (Minneapolis, MN: Twin Cities Innovation Alliance, 2021).

Organizations advocating for increased opportunities for community to be engaged in decisionmaking around data—or to fight data use if community is not involved or the harm is too great—include [Data for Black Lives](#), [Allied Media Projects](#) (supporters of [Our Data Bodies](#) and [A People's Guide to AI](#)), and

participants in the [Civil Rights Table on Data, Privacy and Technology](#). Other organizations like [AI Now Institute](#), [Data & Society](#), and [Upturn](#) are focused on the use of automated decisionmaking systems more specifically.

Concern #3: People who could use data to make personal decisions, improve programs and services, or advocate for effective and equitable government do not have data in the ways they need.

This third failure reflects frustration that the promise of data to improve outcomes for people and communities through the types of uses listed in box 1 hasn't been realized. In some cases, agencies are not sharing information altogether or the data shared are not in a form people could use. In other cases, agencies are not collecting the information that people need. The development of data systems has often been driven by funding streams and their reporting requirements, such as No Child Left Behind and state longitudinal data systems. And those reporting requirements can be driven by what's easiest to quantify and can ignore the complexity of the issues. Test scores [alone don't provide actionable information for teachers and principals to change and improve instruction or for parents to advocate for specific school reforms](#). Many pushed back against No Child Left Behind's focus on using [data as a hammer for accountability](#), including tying teacher evaluations to student performance. This disregard for the interests of frontline staff is found in other sectors as well. For example, [electronic health records](#) are designed with a primary focus on

improving billing and insurance claims and not on improving doctors' interaction with and treatment of patients. In the education sector, the Data Quality Campaign has been a leader in acknowledging this failure and promoting the use of ["data as a flashlight."](#) The American Medical Association is promoting [improvements in electronic health records](#) to better serve doctors and improve care.

Beyond education, other initiatives have focused on the public sector's lack of capacity to use data and evidence to deliver results for residents across agencies. Bloomberg's [What Works Cities](#) initiative recognized that local governments were struggling to fulfill data's potential application toward effective public services and began a program of training and certification to improve practice across the country. [The GovLab](#) also works to strengthen the ability of institutions to collaborate and use data responsibly to solve public problems, such as by providing guidance for [data stewards](#) and championing models of [sharing of private-sector data](#).

Rights and Harms

The failures identified above elevate the importance of acknowledging and safeguarding both individuals and communities (groups of people and places) affected by data use and integration. As people with different perspectives debate over appropriate uses of data, it is helpful to have a common understanding of the rights that all stakeholders should protect and the harms they should prevent in designing approaches to managing and applying data. Many organizations have been promoting individual and collective rights related to data and seeking changes from the government and private sector to protect these rights,

including the [Civil Rights Privacy and Technology Table](#) and [Data for Black Lives](#).

Trust will build over time if data stewards and users are more transparent about how they weigh and balance these rights and steps taken to avoid potential harm. These rights and harms apply to all people and communities, but we recognize that these harms have a disproportionate impact on Black, Indigenous, and people of color and people with limited financial resources.

Rights for Individuals:

- Right to choose how and under what circumstances data about themselves can be shared and used (privacy).
- Right to have data about themselves protected from inappropriate or harmful disclosures (confidentiality).
- Right to view, access, and correct personal data that are collected and shared.

Harms to Individuals:

- Data may be used against an individual to disadvantage them, victimize them through predatory schemes, or stigmatize them in personal or professional settings.
- Data could be shared inappropriately or without context or proper training, which could affect how a person is treated and served.
- Data breaches with personally identifiable information or reidentification could result in identity theft or other harms to individuals.
- Individuals without free and straightforward access to their data have less ability to advocate for themselves and to correct the record.
- Individuals are forced to make decisions about programs or services without access to information they need.

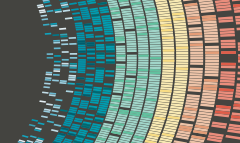
Rights for the Collective:

- Right of marginalized groups, who are more vulnerable to harm from data use, to be prioritized in weighing benefits and risks of data collection, integration, and use.
- Right to have data disaggregated (by race and ethnicity, gender, age, and so on) and contextualized to uncover driving factors, determine differential impacts, and target resources among groups.
- Right to data in formats people can easily digest on programs, services, and their community, especially for groups with the least resources and connectivity.
- Right to access data on government-funded program activities and outcomes for public accountability.

Harms to the Collective:

- People can use data and technology to disproportionately harm people and communities of color, whether through increased surveillance, disinvestment, or predatory schemes.
- Inputs and outcomes of algorithms in decisionmaking reflect biases in our current racist systems and harm people and communities of color.
- Presenting data about groups and communities, particularly without examination of root causes, can perpetuate negative stereotypes and white privilege.
- Structural and institutional racism is not brought to light or addressed, and inequitable policies are reinforced.
- Data are not available to identify needs or inform organizations' and communities' decisions and advocacy.
- Ineffective or harmful public programs are not improved or changed, and effective public programs and policies are not advocated for or developed.

Approaches to Building Trust



To use data to benefit communities, protect the rights of individuals and the collective, and prevent further harms, data stewards and users need to build the trust of the people represented in the data and those affected by its use. No one program or policy will build the necessary trust with a community and engage them to accelerate the use and integration of data as a tool for transformative and sustainable change. Investments are needed at multiple levels to show how data can be collected and used responsibly. We have organized the strategies into four categories, each of which addresses particular rights and harms and implies different types of engagement with communities (table 1).

We briefly describe each category and its approaches, with a discussion of relevant rights and prevention of harms, and the level of engagement with the community. For each approach, we provide a few illustrative examples. Additional examples are included in the appendix A. Although presented separately, we acknowledge that the categories are all connected. For example, laws may set the parameters for governance, or website platforms could facilitate building the capacity of those who want to use the data.

Table 1: Summary of Approaches to Build Trust for Collection and Use of Data

CATEGORY	APPROACH
Enact and refine laws and regulations related to data	<ul style="list-style-type: none">■ Set public policy related to data privacy, sharing, and use that reflects the current environment and latest best practices
Apply technical solutions to expand and control data access	<ul style="list-style-type: none">■ Improve access to one's own data■ Improve access to summary information on programs and services■ Rebalance risks and benefits in research and analysis
Increase community data capacity	<ul style="list-style-type: none">■ Increase awareness and skills of people represented in the data to give them more influence and power■ Build capacity of people in public agencies and other organizations to ethically collect, manage, and use data■ Augment agencies' and organizations' data capacity with external expertise
Establish and enhance governance for data and data systems	<ul style="list-style-type: none">■ Establish processes to ensure security and engagement and that people's information needs are met■ Improve systems for consent and transparency

Enact and Refine Laws and Regulations Related to Data

Refined laws and regulations covering data use and integration by public agencies will provide a foundation for establishing trust with the community if transparency and privacy are prioritized. These laws and regulations should clearly describe how data will be used and provide assurances that data will be protected.

Approach: Set public policy related to data privacy, sharing, and use that reflects the current environment and latest best practices

The purpose of this approach is to codify individuals' rights, require transparency, set parameters for data sharing, and establish collective rights to information. In some cases, new legislation will be required to create or modify policy, but in many cases, these changes can be accomplished via agency regulations.

Following the dissolution of inBloom, there has been progress on recognizing and protecting individual rights to privacy, particularly in the education sector. Hundreds of [education-related privacy laws](#) have been passed since, including [Georgia's 2014 Student Data Privacy, Accessibility and Transparency Act](#), which is one of the best and includes provisions such as explanations as to why data are collected, rights for parents to view and obtain student records, and elimination of unnecessary data collection. However, there are still concerns that privacy is not adequately (or comprehensively) addressed. The European Union's General Data Protection Regulation is currently the best comprehensive example of privacy legislation, governing both private-sector companies and government agencies and aligning with the individual rights described above.⁵ California and Virginia have passed consumer privacy laws, but they generally apply only to large private-sector companies.

Though regulating the private sector is not the subject of this paper, we recognize that the public's concerns about how companies' data collection and use influence their views on the public sector.

In the absence of comprehensive federal data privacy legislation, there are proactive steps jurisdictions and agencies can take to improve transparency and establish collective rights to information, the latter of which are less often addressed in public policy. As one example, the [District of Columbia established a data policy](#) that requires the [inventory](#) of all government datasets, prioritizes sharing data with the public and across agencies, and defines data protection standards. Jurisdictions could also set policies to be transparent about their data destruction practices and harm-reduction measures as well as their use of algorithms or require audits of private algorithms for bias (Milner and Traub 2021). Allegheny County, Pennsylvania, developed a decision-support tool to help prioritize supportive housing services. The county provided [details about the tool's methodology](#) and subjected it to an independent ethical review. Additional examples are shared in the approaches to expanding and enhancing governance below.

⁵ For a detailed look at the provisions of the General Data Protection Regulation see: OneTrust DataGuidance and the Future of Privacy Forum, "[Comparing Privacy Laws: GDPR v. CCPA](#)" (Washington, DC: Future of Privacy Forum, 2019).

Apply Technical Solutions to Expand and Control Data Access

This category describes how to use technology and statistical methods to advance the rights of individuals and communities and build trust by providing greater control and more value from data. We divide the category into three approaches based on the type of data involved: 1) improve access to one's own data, 2) improve access to information on programs and services, and 3) improve access to data for research and analysis.

Approach #1: Improve access to one's own data

As noted in the rights section, individuals should be able to easily view, access, and correct data about themselves, but this is often not possible in today's systems. Data owners can use technology to aggregate people's data across sources and design portals so they can easily view data on themselves and control who can access it. By seeing one's own information, individuals can make better-informed decisions about their and their families' education, health, and employment and advocate for themselves with government agencies or social sector organizations. They will also be in a better position to avoid harm if they are aware of the information an educator, potential employer, or landlord sees about them and can correct misinformation. Developers should engage the intended audience through user-centered design to understand what people would like to know and how data should be presented.

One promising example of this approach is a [learning and employment record](#), a digital record that can

be transferred across systems such as education, the military, workforce training and credentialing programs, and employer human resource systems. Similarly, some governments are piloting programs to allow people experiencing homelessness a digital identity storage locker that would include key documents, like their birth certificate or social security card. This could help them access social services without the delay and worry of keeping track of documents. [Baltimore recently began such a pilot](#) that included feedback from members of the Baltimore Continuum of Care's Lived Experience Advisory Committee in the design of the tool. Recognizing the importance of giving people access to their own data, Allegheny County's Department of Human Services developed an innovative portal, AccessMyInfo, to give clients access to the integrated data collected from dozens of sources about themselves or their dependents. The department worked with program participants to develop the system through a community-centered design process.

Approach #2: Improve access to summary information on programs and services

Technology can also facilitate the right to access aggregated data on program outcomes so people can evaluate services' effectiveness, quality, and equity. With consistent and easy-to-use portals to view this information, people and communities can choose which services to use, advocate for improvements of programs and systems, and provide oversight of government services. This approach involves both ensuring that people have the most-helpful information for evaluating effectiveness, quality, and equity and that the technology delivers this information in a useful, accessible way.

[College Scorecard](#) and [TrainingProviderResults.gov](#) provide data on program cost and details, as well as data around graduate completion, placement, and salaries after finishing different college and training programs, to inform people's decisions around higher education. These stand out for informing decisionmaking on an outcome people care about. At the secondary school level, the [REMIQS](#) project aims to measure more holistically school quality and identify schools that promote disproportionately positive outcomes among marginalized groups, such as students of color or those with low incomes. Focusing on disaggregated outcomes could highlight programs and services that close racial or economic

disparities and should be replicated in other places. In a local example, Great MN Schools' [Minneapolis School Finder](#) includes data that reflect engagement

with parents on what information they need to make decisions about where to send their children to school.

Approach #3: Rebalance risks and benefits in research and analysis

Innovations in statistical methods and technology to produce data for research can advance rights in several ways, particularly the individual right to confidentiality and the collective right to have information on public programs. One example of a statistical innovation that helps protect privacy and expand access is synthetic data, wherein a dataset is created by algorithm to have the same properties as actual administrative records. In other cases, new technologies enable queries of single datasets with individual-level records that return aggregate results or secure linkage of multiple datasets without sharing the full underlying data with external users.

The data collaborative [Birth through Eight Strategy For Tulsa](#) held 12 years of data from government, nonprofit, and the private and philanthropic sectors that represented 32 programs and 220,000 individuals, and they piloted a platform developed by Asemio that used cryptographic record linkage to integrate data while keeping individual and organizational

data secure. This enhanced privacy for individuals and organizations and produced insights to improve service delivery (Actionable Intelligence for Social Policy 2020).

Only a very small number of government analysts and select researchers have access to taxpayer data from the Internal Revenue Service (IRS). IRS staff and the Urban Institute collaborated to develop a synthetic database that preserves many of the statistical characteristics of the restricted data without containing any identifiable tax return information.⁶ The vision is that researchers can test their statistical code on the synthetic data and then submit them to run on IRS computers, subject to a review to guarantee that their output does not disclose private information. The US Census Bureau used similar techniques on unemployment insurance records to develop the [LEHD Origin-Destination Employment Statistics](#) block-level public dataset, which describes where people live and work.

Increase Community Data Capacity

This category has three approaches: 1) increase awareness and skills of people represented in the data to give them more influence and power in how data are collected and used; 2) build capacity of people in agencies and organizations to ethically collect, manage, and use the data; and 3) augment agencies' and organizations' data capacity with external expertise. A [community with data capacity](#) is one where people can access and use data to inform efforts to understand and improve outcomes where they live. To achieve this, communities need both the resources to enable data capacity and the skills and practices to support individuals, organizations, and the community collectively. But data capacity generally is not widely held across communities and is concentrated in more powerful institutions that traditionally have held data and managed access. Residents and organizations should have the skills and supports needed to be informed advocates for the protection of individual and collective rights and participation in data system governance.

Approach #1: Increase awareness and skills of people represented in the data to give them more influence and power

This approach to building trust with the community is among the most foundational. Being a good consumer of data and protecting one's rights does not

require the technical skills to analyze data in statistical programs. But people do need to know where they can find data about programs and, more generally,

6 Len Burman, "[TPC Builds A Moog—Or How Synthetic Data Could Transform Policy Research](#)," Tax Vox (blog), July 13, 2020.

about their local conditions and how to interpret that information to be better equipped to make choices for themselves and their families. This is especially critical for people with fewer financial and information resources. They can also learn what questions to ask about how their data are used, to weigh benefits and risks of providing their consent. Finally, people can also advocate against data uses they consider harmful and participate in the governance of data and data systems. As noted earlier, Black, Indigenous and people of color are disproportionately represented in public data systems because of decades of discrimination and structural racism, which have led to increased surveillance and lack of access to opportunities. As this is the case, special attention should be paid to supporting Black, Indigenous and people of color to fully participate in discussions about data integration and use.

Several initiatives have helped people wield more influence and power when interacting with institutions collecting and managing data about them. The [Digital Defense Playbook](#) by Our Data Bodies provides a series of activities to build a deeper understanding of data and how it affects people's lives. Chicago Beyond's [Why Am I Always Being Researched?](#) describes seven inequities in research design and evidence collection and corresponding ways to balance the power dynamics. It is organized to allow community organizations, researchers, and funders to understand their roles in community-based research. The Twin Cities Innovation Alliance seeks to build community self-determination in defining how big data and predictive tools are used through the [Data 4 Public Good program](#), including facilitation of data advisory bodies, curriculum to raise awareness, and legal and technological expertise.

Approach #2: Build capacity of people in public agencies and other organizations to ethically collect, manage, and use data

Building trust and authentically engaging communities will require the professionals involved in the collection, management, and use of data to work in new ways. This includes developing a culture that values responsible data use and centers the rights and needs of people and communities represented in the data in all stages of data collection, storage, and use. To understand and protect individual and collective rights, people need to master technical skills around security and privacy. They also need to learn best practices in developing relationships to authentically engage people affected by data use and to clearly communicate various aspects of data practice to different audiences. This approach should include the full range of people involved from data collection to application, including managers overseeing the governance, analysts linking data and producing statistics, and frontline workers encouraged to use the data to improve services.

[A Toolkit for Centering Racial Equity Throughout Data Integration](#), by Actionable Intelligence for Social

Policy, describes positive and problematic practices for centering racial equity across the stages of the data lifecycle. It includes templates for deciding who is at the table, engaging communities, and identifying root causes. The [Responsible Data for Children initiative](#), a joint effort between UNICEF and The GovLab at New York University, highlights best practices and actionable principles for responsible data management to put children at the center of data activities. It covers the current state and opportunities for institutional, technical, human, and legal and policy infrastructure. The [Providence After School Alliance](#) is an example of linking data to better serve kids and help frontline workers. The out-of-school time providers were able to use the data to enroll children through a centralized process, better target their services, and understand outcomes for the young people. There are also [many organizations](#) seeking to expand the workforce pipeline to get technologists into government service, which could help improve data capacity. Though, to date, many of these efforts have been focused on the federal government.

Approach #3: Augment agencies' and organizations' data capacity with external expertise

In addition to building the data capacity of professionals' ability to manage and use data, additional capacity can be leveraged from outside one's agency or organization. Several specialized nonprofit and private-sector organizations have missions to collaborate with government agencies and nonprofit organizations around data and technology to further the public good. These are often time-limited collaborations around a specific project or initiative. Alongside leveraging outside expertise in the short-term, these programs seek to introduce new ideas and influence the culture of data use at the hosting institutions. When these partnerships are effective and result in beneficial actions for communities, they counter the failure described

earlier, that people do not have the data they need for making decisions and improving programs.

The assistance can take many forms. The [Strategic Data Project](#) at Harvard University's Center for Education Policy Research trains and places fellows for two years in school systems and education nonprofits. Fellows work on a project selected by the agency to improve the culture of data use. The [Clear My Record](#) team at Code for America partners with courts, state governments, and others to design and implement programs to clear people with eligible criminal records automatically and without burdening them. Their services include technical evaluation, guidance on data infrastructure, support for bulk data processing, qualitative and quantitative research and analysis, and software development.

Establish and Enhance Governance for Data and Data Systems

This category has two approaches: 1) establish processes to ensure security, engagement, and people's information needs are met, and 2) improve systems for consent and transparency. Proper data governance⁷ and inclusive governance structures will protect individual and collective rights; balance the risk and benefits of data collection, integration, and use; and build trust with communities.

Approach #1: Establish processes to ensure security and engagement and that people's information needs are met

All organizations and agencies that collect or handle data should engage in proper data governance. Data governance throughout the data life cycle must be designed to protect the individuals in the data and prevent unauthorized access to the data. Doing so honors people's rights to privacy and confidentiality. More intensive levels of engagement with communities are critical for success in this approach.

Involving data users and people represented in the data in the development and review of data governance programs can help ensure people's information needs are better met, whether they are frontline workers or community advocates. In

some cases, including for integrated data systems and projects and for data on people and groups who have been marginalized, it is essential to have the people represented in the data also engaged in data governance. This typically occurs through a formal [governance committee](#), set up to advise on the policies and procedures of the data governance program. [Broward Data Collaborative](#) used a community-participatory action research approach to build a governance structure that includes youth and community advisors for their integrated data system.⁸ Even if the community hasn't been involved from the beginning, there are ways to engage people that can inform some aspects of data use and integration. For

⁷ Data governance is defined as an approach to data management, which is formalized into a set of policies and procedures that address the full life cycle of data, from acquisition to disposal. For more information, see https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Data_Governance_and_Stewardship_0.pdf.

⁸ See also the case study on page 50 of Actionable Intelligence for Social Policy (2020).

example, [Data Assembly in New York City](#) engaged public stakeholders through a series of “mini-publics” to understand their expectations on responsible reuse of data in the COVID-19 pandemic that resulted in a [Responsible Data Re-Use Framework](#). Expanding governance to include people represented in the

data will also likely lead to different data products, potentially including greater data disaggregation. For example, [Data 4 Black Lives](#), among others, pushed for states to publicly report COVID-19 cases and deaths by race, knowing that structural racism put Black people at disproportionate risk.

Approach #2: Improve systems for consent and transparency

People and communities also need to understand why data are collected, how they are being used, and what potential benefits and risks there are to participating in data collection (including data gathered administratively to operate programs and for research). A person's right to privacy is not respected if they are unable to make an informed choice about sharing their information. In the case of inBloom, the failure to engage the parents and other stakeholders and provide clear examples of what data would be shared or integrated and for what purpose helped fuel the opposition to the initiative (Bulger, McCormick, and Pitcan 2017). In New Zealand, the Data Futures Partnership has promoted a [“path to social licence”](#) and developed guidelines for being transparent about data-use practices to build public acceptance for data use and sharing. They believe that with open and transparent data practices, state and local agencies can build acceptance for data use, creating a social license. Greater transparency and clear opportunities for consent can help address some of the failures of data use and integration described above. The state of

Kentucky's KYSTATS program has been a leader in this area. They have a published [data access and use policy](#) as well as a defined [data request](#) process that includes a data dictionary so it is clear what fields are being collected and integrated.

The full potential of using data to improve people's opportunities and benefit communities will only be realized if the field addresses its failures. National philanthropy has a critical role in this effort to help the public and social sectors prioritize trust building in how they use and integrate data. Going forward, the approaches described in this paper can guide a learning and action agenda for philanthropy that raises the profile of these issues in cross-sector conversations. Funders can also help center racial equity and community needs in these conversations and in their own investments related to data use. Thoughtful review by funders of how their actions and investments contribute to building public trust for using data for good is essential to making programs and policies more ethical, effective, and sustainable.

Roles for Philanthropy

Ways to Advance the Field

Many investments by DFC members and other foundations to date have laid the groundwork for future grantmaking and will continue to evolve and have impact on communities. Though addressing an issue through grantmaking is generally foundations' primary strategy, they also have several other levers to influence the field on how data are integrated and used. Foundations, whether as individual institutions, through DFC, or through other funder networks, should pursue four strategies to maximize their impact: **learn, improve practice with grantees, champion ideas, and invest through grantmaking.**

They can **learn** about current practice and emerging ideas in data use and protections at the national and local levels to understand the broader landscape of efforts and uncover new opportunities. This is good advice for any issue area so lessons can be exchanged across sectors and no effort is wasted reinventing the wheel. It is even more essential in the fast-paced field of public and social sector data use, where resources, innovations, and relevant developments emerge continuously.

They can also **improve practice with grantees** related to data, with whom they already have working relationships of exchange. Grantees may be closer to communities and be able to suggest ways to build trust around data for different groups. Grantees may also have innovative practices funders should pay close attention to and learn from to bring back to their colleagues. Funders can share their perspectives and learning with current grantees, offer workshops

and training, and connect grantees to each other and national support organizations to identify common interests. For potential grantees, they can add proposal requirements asking grantees to describe how they will ensure transparency, incorporate community engagement, and include time in their budget for capacity building.

A third way funders can advance the field is by their power to **champion ideas** to influence people in their own organizations and boards and externally. It includes elevating ideas about data and trust with other funders, adding institutional voices to advocacy campaigns for more funding or responsible legislation, or contributing to advisory groups and convenings. This also includes advocating within one's own institution, whether with peer program officers, learning and evaluation staff, or leadership and the board.

Finally, **investments through grantmaking** is the traditional way foundations have advanced the field. The substance of the grant will align with the funder's interest in sectors, actors, perspective on rights and harms, and other factors, but we believe situating the grants within our approaches to building trust could help with understanding how a funder's portfolio relates to improving data and data systems. Grants may support the development of policy and legal recommendations, planning and implementation of technical innovations, building the capacity of a specific place or actors, and testing new models of governance.

Considerations for Investments

Each funder will review potential investments of time and funding with their specific interests and foundation priorities in mind, but they should also review how the data-related investments are contributing (or could contribute) to building trust with communities. Given the long history of data being weaponized against Black, Indigenous, and people of color, it is also necessary to reflect on whether investments are addressing historical harms to them, or if there are existing racial disparities that should shape how your investments are made. Funders can use the following questions drawn from our review to consider how their actions help to strengthen the current gaps and weaknesses in the field.

- Will the program or activity result in increased meaningful participation of people represented in the data or most affected by their use? Will it shift power from the traditional experts and data stewards?
- Will the program or activity put relevant data in the hands of people to make individual decisions,

improve services, or advocate for effective and equitable government?

- Will it increase the ability of agencies to protect privacy and strengthen legal, regulatory, or governance frameworks?
- Will the program or activity advance more equitable outcomes for Black, Indigenous, and people of color?
- What is the scale of activity's impact, whether through federal and state policy, national networks, or widely applicable resources and lessons?
- Will the program or activity break new ground or pilot a new model that has the potential to increase trust for data use?
- Will the program or activity contribute to sustained improvements to how data are used and integrated, whether through smarter laws and regulations, technical infrastructure, shifts in organizational culture and norms, lasting capacity, or inclusive governance structures?

Areas for Investment

We highlight some promising opportunities below for learning and coinvestments of time and funding for furthering responsible data use and integration. We would recommend that funders reflect on the considerations in the previous section for these or any new investments. This paper presents a framework for approaches to strengthen policy and practice and bolster people's trust for how the public and social

sectors use and integrate data for the benefit of themselves and their communities ([table 1](#)). Though all approaches are needed, we suggest that funders accelerate progress and innovation on the latter two categories, which we believe are less mature: increasing community data capacity⁹ and establishing and enhancing governance for data and data systems.

Learning

- **Review your foundations' investments and strategies** within the approaches to building community trust framework, and reflect on how they are meeting the data needs of people and communities, where there are gaps, and how you could partner with other funders.
- **Develop a learning agenda** to scan the latest activities in the field connected to your grantmaking, and exchange ideas with other funders.
- **Develop a set of principles** around how your foundation views data use and integration and building trust with communities to guide decisionmaking and share with your grantees.
- **Convene discussions** within your foundation about how to socialize collective rights and what they might mean for current and future grants.

Improving practice with grantees

- **Organize listening sessions** with your grantees on how data use is viewed, how trust is being fostered, and what supports their community needs to be able to participate in data governance, oversight, and advocacy. Incorporate findings into your grantmaking process and strategy.
- **Review your own grantmaking process** for requirements or guidance on steps to build trust about data use and integration, such as transparency around data collection and use and requirements for sharing findings with the community.
- **Provide support to grantees** to foster a culture of data use that builds trust with the community and informs continuous improvement.¹⁰ These may include training or workshops on data governance, privacy protections, or community engagement.
- **Host exchanges** with grantees and program colleagues about data use cases that may carry unacceptable levels of risk, such as sharing information between social services or schools with law enforcement.

Championing ideas in the field

- **Identify opportunities** for your foundation to support advocacy for federal and state initiatives that prioritize trust for data use.
- **Share the need for building trust** for data use and integration and examples of strategies to do so with other grantmaker associations to influence local and national foundations.

9 For recommendations on ways that philanthropy can increase community data capacity, see Leah Hendey, Kathryn L.S. Pettit, Jake Cowan, and Marcus Gaddy, ["Investing in Data Capacity for Community Change"](#) (Washington, DC: Urban Institute, 2020).

10 See, for example, Hendey et al. (2020) and von Claus Zastrow and Zeke Perez Jr., ["Using State Data Systems to Create an Information Culture in Education"](#) (Denver, CO: Education Commission of the States, 2019).

Grantmaking

- **Invest in testing models** in different contexts for community participation and advocacy for transparency and monitoring of government use of algorithms and automated decisionmaking systems.
- **Partner with select hosts** of integrated data systems or data sharing initiatives covering different sectors to rethink existing governance and include the participation of community and frontline workers, with a strong learning agenda incorporated to develop guidance for the field.
- **Increase funding** for organizations to build long-term capacity to provide ongoing and responsive technical assistance to state and local actors around data privacy, sharing, and governance.
- **Invest in the assessment of the effectiveness of different data-related trainings** to empower community members (personal privacy, algorithms, governance onboarding, and data use for advocacy, for example) to package and scale training for tailored audiences.
- **Extend state- and local-level reach** by funding relevant professional and peer-learning networks and groups that connect to government and nonprofit staff to spread awareness, build customized materials, and enlist new types of organizations like libraries.¹¹
- **Fund pilots** to test what strategies are needed to make data actionable for frontline workers, such as teachers or case managers.

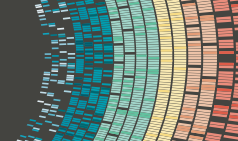
¹¹ See for example, [Civic Switchboard](#), an initiative to develop the capacity of academic and public libraries in civic data ecosystems.

Conclusion

Advances in technology have increased the potential for data use and integration as a tool to address long-standing social issues and disparities in outcomes for children and families. But unless we also make progress on building trust with communities and directly engage them about how and why data are used, data are of limited value. National foundations,

with their positioning and influence, have tremendous opportunities to change the dynamic. They can place communities at the center of their efforts, provide direction to the field, and make investments to pilot and scale innovations in approaches to building trust with communities for data use and integration.

Appendix A



Selected Examples of and Resources for Data Use and Integration

As part of this paper, we collected examples of guidelines, initiatives, and other resources that we or the members of the Data Funders Collaborative considered important to understand the current state of the field. They are generally classified by the primary approach to building trust for collection and use of data they represent, but a few are listed in multiple categories. We did not assess how well these examples center trust or seek to protect rights and prevent harms. Reflecting the interests of many Data Funders Collaborative members, there are more examples from the education field than other fields. It was developed from public information, so we may not have the complete funder list or the year of publishing. The DFC will continue to build on this list of promising approaches and efforts. Please visit the [DFC website](#) for a refreshed version of this list. If you have suggestions for additions to this list, please reach out to DFC Coordinator Aimee Rogstad Guidera at Aimee@GuideraStrategy.com.

Name	Author/Creator	Funder	Year published or started
Enact and refine laws and regulations related to data			
<u>General Data Protection Regulation (GDPR)</u>	European Union	Not applicable	2018
	<i>An example of comprehensive privacy legislation – governing both private sector companies and government agencies and protecting individual rights.</i>		
<u>Foundations for Evidence-based Policymaking Act</u>	U.S. Congress, additional guidance from the White House and Office of Management and Budget	Not applicable	2017
	<i>Legislation that emphasizes collaboration and coordination to advance data and evidence-building functions in the Federal Government by statutorily mandating evidence-building activities, open government data, confidential information protection, and statistical efficiency.</i>		
<u>Principles for Legal Protections for Digital Civil Rights</u>	Civil Rights Privacy and Technology Table	Ford Foundation, and others	2020
	<i>A set of civil rights principles to be upheld in any new legal protections to ensure that technology is designed and used in ways that respect civil rights, preserve privacy, ensure transparency and hold both nation-states and companies accountable for harm.</i>		
<u>Data saves lives: reshaping health and social care with data</u>	UK Department of Health and Social Care	Not applicable	2021
	<i>A draft Digital Transformation plan for the UK to understand the potential for data-driven innovation and improve transparency so the public has control over how the government uses their data; to make appropriate data sharing the norm across health, adult social care and public health to provide the best care possible and to support staff throughout the health and care system; and to build the technical, legal, regulatory foundations to make that possible.</i>		
<u>Georgia's Student Data Privacy, Accessibility and Transparency Act</u>	State of Georgia	Not applicable	2014
	<i>One of the best of the many education-related privacy laws passed after the dissolution of inBloom. It included provisions to explain why data are collected, set out rights for parents to view and obtain student records, and eliminated unnecessary data collection.</i>		

Name	Author/Creator	Funder	Year published or started
A Human Rights-Based Approach to Data	United Nations Office of the High Commissioner for Human Rights	Not applicable	2018
	An approach that helps bring together relevant data stakeholders and develop communities of practice to improve the quality, relevance and use of data and statistics consistent with international human rights norms and principles.		
Data Policy	District of Columbia	Not applicable	2017, 2018
	A city government data policy that requires an inventory of all government datasets, prioritizes sharing data with the public and across agencies, and defines data protection standards.		
Allegheny Housing Assessment	Allegheny County, Pennsylvania	Not applicable	2020
	A decision support tool to help prioritize supportive housing services, for which the county published a detailed methodology and had an independent ethical review of the tool.		
Technical solutions: Improve access to one’s data			
Data backpacks	Digital Learning Now!	Foundation for Excellence in Education	2012
	An electronic student record of data from transcripts and other information that can carry from grade to grade and school to school.		
Data passports	DataPassports (a Software as a Service Technology firm)	Unknown	Unknown
	A data-centric security and privacy solution that enforces privacy and security from end-to-end with transparent protection of data at the source.		
Learning and Employment Records	T3 Innovation Network, managed by the U.S. Chamber of Commerce Foundation	Unknown	Unknown
	A digital record that can be transferred across systems such as education, workforce training and credentials, employer human resource systems, and the military.		
My Digital Storage Locker	City of Baltimore, Amazon Web Services, New America's Digital Impact and Governance Initiative	New America, Kaiser Permanente, Annie E. Casey and Rockefeller Foundations	2021
	A pilot program to allow people experiencing homelessness to have a digital identity storage locker that would include key documents like birth certificates or social security cards. People can control who can access their documents, facilitating access to social services and preventing document loss. Baltimore sought feedback from the Continuum of Care's Lived Experience Advisory Committee.		

Name	Author/Creator	Funder	Year published or started
Human Services Client Portal (see page 24 of AISP Toolkit)	Allegheny County Department of Human Services Client Experience Analytics Unit	Not Applicable	2020
<i>A portal (in development) that integrates school district, housing, income support, child welfare, and other individual-level data, housed in the county's data warehouse and accessible by clients and network providers.</i>			
Technical solutions: Improve access to aggregate information on services, programs			
REMIQS	KnowledgeWorks	Barr Foundation, Carnegie Corporation of New York, the Nellie Mae Education Foundation, and Oak Foundation	2020
<i>This project aims to more holistically measure school quality and identify schools that promote disproportionately positive outcomes among marginalized groups, such as students of color or those with low incomes.</i>			
College Scorecard	U.S.Department of Education	Not applicable	2015
<i>A portal that provides data on salaries that people earn after completing different college programs to give people more information as they choose which one is right for them.</i>			
Great Schools	GreatSchools	Walton Family Foundation, Bloomberg Philanthropies, and multiple others	2015
<i>A portal for parents to see ratings of individual schools in the areas of test scores, student progress, and equity. The ranking system has been critiqued for penalizing schools that serve largely low-income students and those serving largely black and Hispanic students.</i>			
Education Opportunity Project	Stanford University	Institute of Education Sciences, the Spencer, the William T. Grant, the Bill and Melinda Gates, the Overdeck Family, and the Russell Sage foundations.	2016
<i>A national database of academic performance. including average test scores, learning rates, and trends in test scores, to help scholars, policymakers, educators, and parents learn how to improve educational opportunity for all children.</i>			
ParentAware	Minneapolis Department of Human Services	Not applicable	
<i>A portal with tools and resources to help families find the quality childcare and early education programs their children need to succeed in school and life.</i>			

Name	Author/Creator	Funder	Year published or started
Minneapolis School Finder	Great MN Schools	Local foundations and individual supporters	2020
<i>A portal for transparent and parent-friendly school quality information that was developed with parents and community partners. Included are profiles of 154 schools, with information on academic performance and a guide to help find a great school.</i>			
Technical solutions: Re-balance risks and benefits in research and analysis			
Synthetic Internal Revenue Service Data	Urban Institute	Arnold Ventures	2020
<i>A synthetic database that preserves many of the statistical characteristics of the restricted data without containing any identifiable tax return information.</i>			
LEHD Origin-Destination Employment Statistics (LODES)	Census Bureau	Not applicable	2001
<i>Synthetic data created from unemployment insurance records collected from the states. Data describing where people live and work, along with demographic and job characteristics, are made public at the census block level.</i>			
Birth through Eight Strategy For Tulsa (see page 24 of AISP Toolkit)	Multiple organizations	George Kaiser Family Foundation	Unknown
<i>A system with 12 years of data from government, nonprofit, private sector, and philanthropic sectors that represented 32 programs and 220,000 unique individuals. BEST piloted a platform developed by Asemio that used cryptographic record linkage to integrate data while keeping individual and organizational data secure.</i>			
Boston Women's Workforce Council	Unknown	Unknown	2016
<i>A project using secure multi-party computation to measure the earnings gap across Boston employers by demographic (age, race, gender) without a trusted intermediary having to hold the sensitive data.</i>			
NCES Secure Multiparty Computation Demonstration	Georgetown's Massive Data Institute	Alfred P. Sloan Foundation	2021
<i>A demonstration to produce an equivalent to the National Postsecondary Student Aid Study report, which currently requires two groups within NCES to exchange datasets with personally-identifiable information, by linking the data and computing the necessary statistics using secure multiparty computation.</i>			
How Data Governance Technologies Can Democratize Data Sharing for Community Well-being	Cambridge University Press	Not Applicable	2021
<i>An article in Data and Policy arguing that data governance technologies have a large role to play in democratizing data safely and in a trustworthy manner. It provides paths and examples for how technology can overcome the lack of data literacy and skills, navigating privacy laws, auditing data access, and developing evidence of need.</i>			

Name	Author/Creator	Funder	Year published or started
Community data capacity: Increase awareness and skills of people represented in the data			
Digital defense playbook	Our Data Bodies	Unknown	2018
	A workbook with education activities for the public focused on data, surveillance, and community safety to create and share knowledge, analyses, and tools for data justice and data access for equity.		
Why am I always being researched?	Chicago Beyond	Chicago Beyond	2019
	An equity-based approach to research to shift the power dynamic and restore communities as authors and owners. It presents the inequities in current research practices and then offers sections for community organizations, researchers, and funders.		
Data 4 Public Good program	Twin Cities Innovation Alliance	Ford Foundation, McKnight Foundation, and Communities for Just Schools	2020
	A program that seeks to build community self-determination in defining how big data and predictive tools are used, including facilitation of data advisory bodies, curriculum to raise awareness, and legal and technological expertise.		
Investing in Data Capacity for Community Change	Urban Institute	Robert Wood Johnson Foundation	2020
	A brief encouraging foundations to invest in community data capacity by promoting the use of data by grantees, championing community data resources and capacity, and investing directly in building community-wide capacity. It also presents a framework for understanding the elements of community data capacity .		
National Neighborhood Indicators Partnership	Urban Institute, local partners	Annie E. Casey Foundation and others	1996
	A learning network coordinated by the Urban Institute that connects partner organizations in 30 cities that ensure their communities have access to data and the skills to use information to advance equity and well-being across neighborhoods.		
Community data capacity: Build capacity of people in agencies and organizations			
Why am I always being researched?	Chicago Beyond	Chicago Beyond	2019
	An equity-based approach to research to shift the power dynamic and restore communities as authors and owners. It presents the inequities in current research practices and then offers sections for community organizations, researchers, and funders.		
A Toolkit for Centering Racial Equity Throughout Data Integration	Actionable Intelligence for Social Policy	Annie E. Casey Foundation	2020
	A toolkit focused on data sharing and integration describing positive and problematic practices for centering racial equity across the data life cycle. It offers three activity templates for deciding who should be at the table, mapping assets and engaging community, and identifying root causes through factor analysis.		

Name	Author/Creator	Funder	Year published or started
National Neighborhood Indicators Partnership	Urban Institute, local partners	Annie E. Casey Foundation and others	1996
	A learning network coordinated by the Urban Institute that connects partner organizations in 30 cities that ensure their communities have access to data and the skills to use information to advance equity and well-being across neighborhoods.		
Responsible Data for Children Initiative	UNICEF and GovLab	Unknown	2019
	A joint effort between UNICEF and The GovLab at New York University that highlights best practices and actionable principles for responsible data management to put children at the center of data activities. It covers the current state and opportunities for institutional, technical, human, and legal and policy infrastructure.		
Student-Centered Learning Research Collaborative	Students at the Center	Nellie Mae Education Foundation and Jobs for the Future	2019
	An RFP designed by Student-Centered Learning Research Collaborative to support projects with researcher positionality integrated into the proposals.		
Principles for Equitable Data Practice	Urban Institute	Robert Wood Johnson Foundation	2020
	A brief presenting principles for equitable data practice across the data life cycle based on the principles of beneficence, respect for persons, and justice.		
Investing in Data Capacity for Community Change	Urban Institute	Robert Wood Johnson Foundation	2020
	A brief encouraging foundations to invest in community data capacity by promoting the use of data by grantees, championing community data resources and capacity, and investing directly in building community-wide capacity. It also presents a framework for understanding the elements of community data capacity .		
Student Privacy Compass	Future for Privacy Forum	Unknown	2014
	A website for finding information, news, and opinions on maintaining student data privacy.		
Data Responsibility Journey	GovLab	Unknown	2021
	An assessment tool that outlines the opportunities and risks to consider at each stage of the data lifecycle when implementing a data collaborative.		
Providence After School Alliance	Providence After School Alliance	Mix of public and private funds	Unknown
	An Alliance that used an integrated data system of youth for out of school time service providers in Providence to inform the development of its AfterZone model.		
Community data capacity: Augment agencies’ and organizations’ data capacity			
DataKind DataCorps	DataKind	Rockefeller Foundation, Microsoft, and others	2015
	A program that pairs “pro bono data scientists with social change organizations on long-term projects that use data science to transform their work and their sector.”		

Name	Author/Creator	Funder	Year published or started
Clear My Record	Code for America	The Ballmer Group	2018
	A partnership with courts, state governments, and others to design and implement programs to clear people with eligible criminal records automatically and without burdening them. Their services include technical evaluation, guidance on data infrastructure, support for bulk data processing, qualitative and quantitative research and analysis, and software development.		
Data Science for Good	Carnegie Mellon University	Unknown	2018
	A full-time summer program to train aspiring data scientists to work on equitable machine learning, data science, and artificial intelligence projects on real-world problems with social impact in partnership with governments and nonprofits.		
Strategic Data Project	Harvard University's Center for Education Policy Research	Bill & Melinda Gates Foundation and Broad Foundation	2012
	A fellowship program that trains and places analytic talent for two years in school systems and education nonprofits. They also have an Institute for Leadership in Analytics that offers practitioners virtual modules to learn about various aspects of analyzing and using data.		
Data4Change Accelerator	Cloudera Foundation	Unknown	2020
	An accelerator to equip nonprofit organizations with the tools and expertise needed to test whether they can generate added value from their data.		
Thriving Earth Exchange	American Geophysical Union (AGU)	Unknown	Unknown
	An exchange that matches community leaders with vetted volunteer scientists for individual consultations, project design, and management services, all pro bono.		
Governance: Establish processes to ensure security, engagement, and that people's information needs are met			
Broward Data Collaborative	Children's Services Council of Broward County	Not Applicable	2018
	A Collaborative that used a community-participatory action research approach to build a governance structure that includes youth and community advisors for their integrated data system.		
Responsible Data Re-Use Framework	GovLab	Henry Luce Foundation	2020
	A framework that resulted from GovLab's The Data Assembly engagement with public stakeholders through a series of "mini-publics" to understand their expectations on responsible re-use of data in the COVID-19 pandemic.		
A Toolkit for Centering Racial Equity Throughout Data Integration	Actionable Intelligence for Social Policy	Annie E. Casey Foundation	2020
	A toolkit focused on data sharing and integration describing positive and problematic practices for centering racial equity across the data life cycle. It offers three activity templates for deciding who should be at the table, mapping assets and engaging community, and identifying root causes through factor analysis.		

Name	Author/Creator	Funder	Year published or started
Designing a Process for Inclusive Innovation: A Radical Commitment to Equity	Digital Promise	Bill & Melinda Gates Foundation	2019
	A paper that presents <i>Inclusive Innovation</i> , a model that reimagines authority, decision-making, and risk in the context of education R&D and provides an overarching framework for authentically engaging underrepresented stakeholders at the earliest stages and shifting their roles to leaders, participants, and beneficiaries.		
Data 4 Black Lives Policy working group	Data 4 Black Lives	Unknown	2020
	A working group that explores alternative models of data governance. The policy working group creates a space for movement builders, data scientists, public interest technologists, legal scholars, and people whose lives are directly impacted by technological bias to collaborate and identify approaches and risks to data governance frameworks.		
Sharing Data for Social Impact: Guidebook to Establishing Responsible Governance Practices	Georgetown's Beeck Center for Social Impact and Innovation	Rockefeller Foundation	2020
	A report that lays out a three-phased approach to adopting sustainable data sharing governance practices: build the collective, define operations; and drive impact. It includes resources, case studies, and best practices.		
Roadmap for Cross-Agency Data Governance	Data Quality Campaign	Bill & Melinda Gates Foundation and Michael & Susan Dell Foundation	2018
	A roadmap with recommendations for states working on developing and implementing a high-quality data governance committee across agencies, such as the state education agency and early childhood, higher education, and workforce agencies.		
Data Governance for Two-Generation Programs	Annie E. Casey Foundation	Annie E. Casey Foundation	2019
	A toolkit designed for two-generation initiatives to think through and refine their data governance practices.		
Participatory Data Stewardship	Ada Lovelace Institute	Nuffield Foundation	2021
	This report proposes a 'framework for participatory data stewardship', using the 'ladder of citizen participation' to explore different means of participation in data governance and increase people's power and agency in decisions about their data.		
Governance: Improve systems for consent and transparency			
Principles for Legal Protections for Digital Civil Rights	Civil Rights Privacy and Technology Table	Ford Foundation, and others	2020
	A set of civil rights principles to be upheld in any new legal protections to ensure that technology is designed and used in ways that respect civil rights, preserve privacy, ensure transparency and hold both nation-states and companies accountable for harm.		

Name	Author/Creator	Funder	Year published or started
<u>A Path to Social Licence</u>	Data Futures Partnership	New Zealand Government	2017
	<i>Guidelines for being transparent about data use practices to build public acceptance for data use and sharing. The partnership believes that with open and transparent data practices, state and local agencies can build acceptance for data use - creating a social license.</i>		
<u>KYSTATS</u>	State of Kentucky	Not applicable	Unknown
	<i>A state data access and use policy and with a defined <u>data request</u> process that includes a data dictionary so it is clear what fields are being collected and integrated. The state has been a leader in providing greater transparency on the data they collect and how it gets used.</i>		
<u>Seattle Privacy and Data Statement</u>	City of Seattle	Not applicable	Unknown
	<i>A web page that provides residents with the city's <u>privacy principles</u> and links to pages that describe the <u>types of data</u> the city agencies typically collect and the ways the city uses information.</i>		
<u>First Do No Harm: Ethical Guidelines for Applying Predictive Tools Within Human Services</u>	MetroLab	Annie E. Casey Foundation	2017
	<i>A report with brief, practical guidance to human service agency leaders on how they can mitigate the risks that come with using predictive analytics. It is structured around four principles—engagement, model validation, review and transparency.</i>		

References

Actionable Intelligence for Social Policy. 2020. [A Toolkit for Centering Racial Equity Throughout Data Integration](#). Philadelphia: University of Pennsylvania.

Bulger, Monica, Patrick McCormick, and Mikaela Pitcan. 2017. [“The Legacy of inBloom.”](#) New York City: Data & Society Research Institute.

Milner, Yeshimabeit and Amy Traub. 2021. [Data Capitalism + Algorithmic Racism](#). New York City: Demos and Data for Black Lives.

Moss, Emanuel and Jacob Metcalf. 2020. [Ethics Owners: A New Model of Organizational Responsibility in Data-Driven Technology Companies](#). New York City: Data & Society Research Institute.

Mozilla Insights, Jonathan van Geuns, and Ana Brandusescu. 2020. [What Does it Mean? Shifting Power Through Data Governance](#). San Francisco: Mozilla Foundation.

Sparapani, Tim and Justin Sherman. 2021. [Privacy Tech's Third Generation: A Review of the Emerging Privacy Tech Sector](#). Washington, DC: Future of Privacy Forum.