

SHARING DATA *for* SOCIAL IMPACT: GUIDEBOOK *to* ESTABLISHING RESPONSIBLE GOVERNANCE PRACTICES

*A guidebook produced by
Natalie Evans Harris,
program fellow with the*

beekcenter
social impact + innovation

ABOUT THE BEECK CENTER

[The Beeck Center](#) The Beeck Center works alongside experts and students to surface, accelerate, and scale promising social impact efforts that drive institutional-level change. Positioned as a neutral, academic-based player, the Beeck Center has built credibility working across sectors to generate new ideas that will provoke leaders to think and act differently.

ABOUT DIGITAL SERVICES COLLABORATIVE

This project is part of the [Digital Service Collaborative](#), a partnership between the Beeck Center and The Rockefeller Foundation that leverages the network of professionals working on data and digital services to scale solutions for greater impact. The Digital Service Collaborative brings together members of the data and digital service community to work together solving specific problems and sharing those solutions throughout the network.

ABOUT THE AUTHOR

[Natalie Evans Harris](#) has spent nearly twenty years advancing the public sector's strategic use of data, including a sixteen year career at the National Security Agency, and eighteen months with the Obama Administration. She co-founded and currently serves as Head of Strategic Initiatives of [BrightHive](#), a data trust platform to help organizations, networks, and communities securely and responsibly link their data to enhance their impact, empower individual and collective decision making, and increase equity of opportunity. She founded the [Community-driven Principles for Ethical Data Sharing \(CPEDS\)](#) community of practice focused on strengthening ethical practices in the data science community through crowd-sourcing a Data Science Code of Ethics. As a Senior Policy Advisor to the US Chief Technology Officer in the Obama Administration, Natalie founded the [The Data Cabinet](#) - a federal data science community of practice with over 200 active members across more than 40 federal agencies.

ACKNOWLEDGEMENTS

I would like to thank [The Rockefeller Foundation](#) for providing the financial support that made this work possible. I would also like to thank the Beeck Center team for supporting me through this year plus journey. My deepest appreciation goes to the Student Analysts who conducted research and contributed to the creation of this Guidebook: [Nitya Biyani](#), [Moonlan Zhang](#), [Robert Roussel](#), and [Ben Lang](#). For the experts that contributed time for interviews, emails, phone calls, and reviewing the many iterations of this report. In particular, I am forever in-debt to: [Bill Yock](#), [Sara Holoubek](#), [Kathryn Pettit](#), [Bill Hoffman](#), [Reg Leichty](#), [Leah Hendey](#), [Michelle Shevin](#), [David Park](#), [Dan Wu](#), [Kristin Wolff](#), and [Kevin O'Neil](#). Finally, I must thank my colleagues at [BrightHive](#) for showing me everyday that the practices recommended in this guide are not only doable but impactful.



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).

THE GUIDEBOOK: HOW TO USE IT

Who is this guidebook for?

This guidebook is for policy makers, whether elected or appointed, and policy staff looking to understand how to leverage data and data sharing towards evidence-based policy making. It can be used for policy makers and organizations interested in giving agency to individuals to allow for full consent to the use of their data, along with organizations interested in ethically and responsibly sharing data. This guidebook will help actors drive impact with stakeholders.

ACTORS

Those who take action on the data

Those who use data to drive impact

Examples: policy makers; federal, state and local staff; school districts

STAKEHOLDERS

Those with a vested interest in data sharing

Those from whom data is being collected

Examples: parents; community members; program managers; IT staff

What will you find in this guidebook?

This guidebook is essentially a resourced framework. It enumerates specific phases, steps, considerations, and outcomes that creators of data-sharing partnerships for social impact should consider. It includes resources throughout and additional resources for a deeper dive in the [annotated appendix](#).

Why was this guidebook created?

The adoption of data-sharing governance practices has traditionally been limited and under-resourced, with the sustainability and effectiveness of these practices often harmed by leadership, capacity, and strategies that are not forward-thinking. To match a growing demand for innovative and data-driven approaches to achieving social impact, we need to reform data-sharing governance practices. Many resources exist to help stakeholders establish processes for data sharing. These resources often demand capacity that stakeholders lack, or limit stakeholders' flexibility. As a result, projects can be overly ambitious or lack specificity or impact.

What Is this guidebook NOT useful for?

Written resources alone may not be enough to generate buy-in from leadership, data teams, partners, or funders. Creating thoughtful and forward-thinking data-sharing partnerships for social impact still has significant barriers that only motivated leaders and good timing can overcome. Getting leaders to invest significant resources at the start of a project may ensure responsible and sustainable operations, but this can sometimes come at the expense of quick wins, having results to show early on. Stakeholders who appreciate new approaches to creating social impact is a critical precursor to impactful data-sharing innovations.

What are some quick tips on how to use this guidebook?

Start by reading the summary and the table of contents for an overview. If you want to create a collective from the ground up, this overview will help you identify your biggest opportunities and your most significant barriers, as well as think through who you should be working with inside and outside of your organization. If you are part of a collective—a group of three or more organizations with a shared goal for driving impact through linked data—or designing one and want to make sure you are on an effective and sustainable track to generating social impact, you might want to start from the beginning of this guidebook. Consider consulting the resources listed in the appendix, rethinking your stakeholder mix, reconfiguring your team, and/or redesigning your approach based on the insights generated by this guidebook.

EXECUTIVE SUMMARY

Many government agencies have well-established use cases for sharing data in pursuit of improved social service delivery in areas such as [K-12 education](#), public transportation, and [healthcare](#). Florida's Broward County, for example, instituted a community feedback loop on data sharing to strengthen family incomes. [Los Angeles County](#) created an integrated data system (IDS) for the delivery of services to its homeless population with a template that outlines data-use practices and procedures for each agency with a form that describes specific data-use policies for each project, the details of the project, and the associated list of data sources for agency approval. On a larger scale, the National Committee on Vital and Health Statistics held a roundtable about self improvement and building trust in the community to drive impact.

But cross-sector data-sharing collaboratives have been a more recent phenomenon, which private companies, governments, and nonprofits are increasingly forming them in support of the social good. While we've seen many success stories in deploying data-sharing collaboratives, many government agencies are also experiencing the pains of data-sharing governance practices that are limited and under-resourced, resulting in practices that can be unsustainable, not effective, and not forward-thinking.

Organizational rules for data sharing typically advance risk mitigation strategies focused on restricting the availability of data to external actors. In an environment where data are only used for making funding decisions or to narrowly evaluate programs, this model can work well. In pursuit of innovation or improved social service delivery, this model is less encouraging. As the amount of data and methods for collecting it increase, so have opportunities for drawing insights about society. Bringing together diverse data sources is crucial to ensuring that insights promote equitable growth. And as promising as data sharing is for improving societal outcomes, the analysis of integrated data (especially through predictive analytics) can easily replicate inequities learned from patterns of past service delivery. Contextualizing data analysis with methods used by the social sciences and ongoing community engagement is crucial to ensuring data analytics do not replicate or worsen inequitable outcomes.

There is a fundamental need to reform data-sharing governance practices to ensure their effectiveness and sustainability. We advance here a guidebook, each step of which is equipped with resources and case studies, to aid stakeholders and actors as they navigate the legal, technical, organizational, and societal challenges to creating good governance practices for responsible and impactful data-sharing collaboratives. This framework aims to provide a holistic guide that encourages data-sharing parties to consider technical, legal, as well as ethical components of sharing. It is divided into three main phases:

Phase 1: Build the Collective

- Understand gap filled by data sharing
- Understand data being shared
- Understand ethical implications

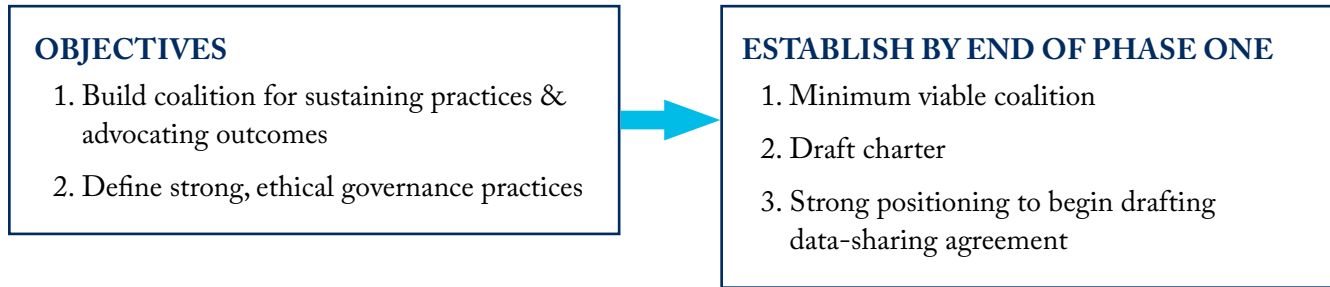
Phase 2: Defining the Operations

- Define success
- Establish internal trust within the collective
- Define operations of the collective

Phase 3: Driving Impact

- Formalize ethical principles and guidelines
- Establish feedback loop for evaluating impact metrics
- Create internal monitoring and assessment

Build the collective: get everyone on board. Start with the policy problem. Identify stakeholders. Take stock of capacity, motivations, barriers, and potential data solutions. Demonstrate value and reduce uncertainty to generate buy-in. Establish a minimum viable coalition and enshrine your shared vision in a charter.



Define the operations: get everyone in line. Create the governance framework tied to the charter. Design a feedback loop and integrate it into the governance framework. Formalize those two elements into a data-sharing agreement. Launch the operations of the minimum viable coalition.

BY ALL PARTIES

- 1. Define governance activities tied to the charter and to overcoming barriers identified in Phase One
- 2. Sign data-sharing agreement to formalize governance structure
- 3. Determine how agreement will be implemented by governing board

BY THE BOARD

- 1. Put into practice governance framework that defines the roles & responsibilities of each stakeholder
- 2. Put into place funding & decision-making mechanisms
- 3. Put into place process for entering & leaving collective

BY YOU

- 1. Formalize mechanism for evaluation & improvement with impact metrics that agree with your theory of change
- 2. Enumerate ethical principles, connected to specific actions, baked into implementation & review process

Drive impact: get everyone to improve and share. Re-evaluate assumptions, approach, and metrics. Survey impacted communities and stakeholders. Use feedback loops to enact iterative improvements to the governance structure. Repeat this process until feedback becomes minimal. Scale up. We recognize that many different actors will be involved in this process and that each actor may face unique challenges, goals, motivations, and opportunities. This guidebook is for people looking to understand how they can leverage data and data sharing towards evidence-based policy making. Moreover, it can be used by policy makers and organizations interested in giving agency to individuals over their data along with organizations interested in ethically and responsibly sharing data.

Two Main Elements of Driving Impact

Element one: Creating iterative improvement

Goal: Gather feedback & incorporate lessons into the data-sharing governance and the agreement that formalizes it, formed in Phase Two

Element two: Sharing best practices

Goal: Model actions after those of other successful endeavors & share what works for an organization like yours with others going forward



GEORGETOWN UNIVERSITY