

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.

DATA SHARING TODAY

Resources

[Understanding Corporate Data Sharing Decisions \(Future of Privacy Forum\)](#)

This report covers the Future of Privacy Forum's research on how companies make data available, why companies share data, and the risks perceived with sharing.

[The U.S. Data Federation Wants to Make It Easier to Collect, Combine, and Exchange Data across Government \(General Services Administration\)](#)

A paper from 18F discussing the potential new ease of data sharing and exchange across government through the GSA Technology Transformation Systems.

[AISP Network Site Case Studies \(Actionable Intelligence for Social Policy\)](#)

A group of case studies on securing and maintaining legal agreements, establishing and adapt governance processes, managing data and analytic processes, and confronting economic and political realities in sustaining Integrated Data Systems (IDS) operations, documenting the group's own exemplary research uses.

[Sharing Data is a Form of Corporate Social Responsibility \(Harvard Business Review\)](#)

An article outlining a new thought process for public-private data sharing.

[Personal Data for the Public Good: New Opportunities to Enrich Understanding of Individual and Population Health \(Health Data Exploration Project\)](#)

A report on enhancing how to generate knowledge and outcomes from health data concentrating on the opportunities of privacy and data ownership, informed consent, data sharing and access and data quality.

[Clinical Trial Participants' Views of the Risks and Benefits of Data Sharing \(New England Journal of Medicine\)](#)

In this article, Mello et al. track opinions on data sharing, concluding that few trial participants had fears about sharing their information.

[It's Time for Data Ethics Conversations at Your Dinner Table \(Tech at Bloomberg\)](#)

An opinion piece that emphasizes the importance of having conversations on data.

[Data Driven Social Partnerships: Exploring an Emergent Trend in Search of Research Challenges and Questions \(Government Information Quarterly\)](#)

In this article, Susha, Grönlund, and VanTulder emphasize the benefit of data collection for social impact and the common good.

[The New Ecosystem of Trust \(Nesta\)](#)

This blog post builds on a series of Nesta think pieces on data and knowledge commons published over the last decade and current practical projects that explore how data can be mobilised to improve healthcare, policing, the jobs market, and education.

[Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information \(Pew Research Center\)](#)

An article summarizing recent research that discusses Americans' opinions about the privacy of data, their lack of control over privacy, and their overall feelings about data sharing and collection.

[Data Governance \(Colorado Department of Education\)](#)

A Phase Three exemplar of how to create a governance board for a community.

[Pathways to Grade-Level Reading \(North Carolina Early Childhood Foundation\)](#)

A Phase Three exemplar that emphasizes driving impact through data collection (and subsequent action).

[Supporting Community Data Engagement—An NCVHS Roundtable \(National Committee on Vital and Health Statistics\)](#)

A Phase Three exemplar that held a roundtable on effectively engaging the community

[Empowering Families Project \(Children's Services Council of Broward County\)](#)

A Phase Three exemplar of a family-centered project to improve educational outcomes through an integrated data repository that included important items such as a community feedback loop to continually improve processes.

PHASE ONE: BUILD THE COLLECTIVE

1.1 Define the Collective

Considerations

- *What problem are you solving?*
- *What impact are you pursuing and what is your theory of change?*
- *Who are the key stakeholders?*

Resources

[Are You Solving the Right Problems? \(Harvard Business Review\)](#)

An article that talks about the importance of defining a problem and how an organization might go about ensuring that it is solving the right one.

[Defining and Analyzing the Problem \(University of Kansas\)](#)

Chapter 17 of the Community Tool Box focused on helping you think through what the problem you are working on is and whether you should pursue it.

[Problem Definition Worksheet \(Wageningen University & Research\)](#)

A worksheet in the Multi-Stakeholder Partnerships portal to help you define your problem in depth.

[All In \(Data for Community Health Online\)](#)

An online community of individuals dedicated to improving health through multi-sector data sharing and collaboration, designed to help you connect to professionals tackling common challenges, share resources and news, and learn about new ideas and best practices.

[The 100 Questions Initiative \(GovLab\)](#)

This initiative seeks to map the world's most pressing, high-impact questions that could be answered if relevant datasets were leveraged in a responsible manner.

[IssueLab \(Candid\)](#)

One of the largest open repositories for social sector knowledge; IssueLab helps organizations better manage, synthesize, and share what they are learning.

[Social Value Proposition \(Forbes\)](#)

A four step framework on building social value propositions that are compelling and relevant.

[Theory of Change: A Practical Tool for Action, Results and Learning \(AECF\)](#)

This 49-page report will help you understand how to create a Theory of Change and align it with your community for optimal results.

[Matching Administrative Data to Inform Policy \(Northwestern University\)](#)

Newsletter article describes an effort at the university to create a new network, set goals for the network, and start to build community around matching administrative data for policy evaluation.

[Five Principles for Applying Data Science for Social Good \(O'Reilly\)](#)

A blog post on best practice principles for using big data for the social good, written by Jake Porway from DataKind. The second principle discusses the difficulty of finding problems. The fourth principle says that teams need a diverse set of stakeholders because they need a diverse set of viewpoints.

[AISP Network Site Case Studies \(Actionable Intelligence for Social Policy\)](#)

The importance of a top-down, problem-first approach to getting an organization on board is covered in several Integrate Data System (IDS) case studies. For example:

- a) [A study](#) of how New York City set up its IDS found that programmatic needs have to come first and drive all decision-making, or else partners will not see the value in sharing.
- b) [A study](#) of how Los Angeles County developed its IDS found that focusing too much on the technology at first was a mistake, and that what really allowed success was to find one person at the top who can cheerlead a project into existence.
- c) [A study](#) of how Washington State's DHHS developed its IDS found that getting leadership to generate momentum was key, so that the value of sharing was recognized, capacity was generated, and initial reluctance to share could be overcome.

[Learning in Action: Evaluating Collective Impact \(Stanford Social Innovation Review\)](#)

An article that will help you think through what collective impact is and why it is essential to consider in the context of partnerships for tackling social problems.

[Stakeholder Analysis Guidelines \(World Health Organization\)](#)

A detailed report from Partnerships for Health Reform on how to precisely define your policies and stakeholders.

[Stakeholder Analysis: Securing the Buy-In You Need \(Lucidchart\)](#)

A blog article on how to identify your stakeholders and get their buy-in.

[The Three Goals and Five Functions of Data Stewards \(GovLab\)](#)

A blog post describing what a data steward is and what roles and goals data stewards have in pursuing social impact in the information era.

[Scaling Big Data for Social Good: The Need for Sustainable Business Models \(GSMA\)](#)

A report on making sure that your social impact solution is sustainable and will scale up.

[Using Data for Action and for Impact \(Stanford Social Innovation Review\)](#)

This article reminds readers that the nonprofit and public sector have to deal with many more stakeholders and outcomes than private companies and for this reason they often struggle with how to properly acquire and use data responsibly.

1.2 Establish Shared Vision

Considerations

- *What are the motivations and goals of stakeholders in wanting to share data?*
- *What are the barriers to sharing data and can they be overcome?*

Resources

[Data Collaboration for the Common Good \(World Economic Forum\)](#)

See page 12 of section 2, part 1 on entitled “Stakeholder Alignment” that gives tips for building trust through stakeholder alignment, advice for responding to challenges and tools and resources for these issues.

[Announcing the Data Collaboratives Research Network \(GovLab\)](#)

A blog post announcing and offering basic descriptions of the data collaboratives effort by GovLab at New York University.

Resources (barriers generally)

[Understanding Corporate Data Sharing Decisions \(Future of Privacy Forum\)](#)

A paper that explores when and why corporations donate data for academic research.

[The Potential of Social Media Intelligence to Improve People’s Lives \(GovLab\)](#)

This resource lists examples and a detailed discussion for each of the five kinds of motivation.

[Designing and Implementing Cross-Sector Collaborations: Needed and Challenging](#) (*Public Administration Review*)

A well-written academic paper from Bryson, Crosy, and Stone on the barriers to cross-sector collaboration.

[Data Driven Social Partnerships: Exploring an Emergent Trend in Search of Research: Challenges and Questions](#) (*Government Information Quarterly*)

A detailed literature review from Susha, Grönlund, and Van Tulder sketching out the field of data sharing for social good and detailing what barriers and solutions others have found.

[Barriers to Using Government Data](#) (Bipartisan Policy Center)

An analysis of findings from the U.S. Commission on Evidence-Based Policymaking's Survey of Federal Agencies and Offices that details the context, barriers, and next steps to sharing government data.

[A Solutions-Based Approach to Building Data-Sharing Partnerships](#) (*Journal for Electronic Health Data and Methods*)

A paper by Wieche et al. describing a taxonomy of barriers and solutions to sharing data for health research.

[A Systematic Review of Barriers to Data Sharing in Public Health](#) (*BMC Public Health*)

A literature review by van Panhuis et al. describing a taxonomy of barriers and solutions to sharing data for health research.

[A Decision Model for Data Sharing](#) (13th International Conference on Electronic Government)

A literature review of data-sharing governance practices from Eckartz, Hofman, and Veenstra that includes three brief case studies in the logistics sector and a data-sharing decision model to help actors identify and address barriers to sharing.

[Lessons Learned in Using Hospital Discharge Data for State and National Public Health Surveillance](#) (*Journal of Public Health Management and Practice*)

A case study from Love, Rudolph, and Shah of the strengths and barriers to sharing health data in a collaboration between federal and state government.

[The Law, Politics and Ethics of Cell Phone Data Analytics](#) (Data Pop Alliance)

A framework, along with a discussion of the legal, political, and ethical implications, for sharing mobile data for the social good.

Resources (legal barriers)

[Navigating Legal Parameters for Cross-Sector Data Collaboration](#) (National Center for Complex Health and Social Needs)

A brief on the legal issues of sharing data in the health sector that yields surprisingly generalizable lessons.

[Legal Issues for IDS Use: Finding a Way Forward \(Actionable Intelligence for Social Policy\)](#)

A report that outlines some findings from an expert panel on setting up a cross-agency integrated data system (IDS), including common concerns, tools and tips, specific laws that could create legal issues, and potential legal instruments that could be used as solutions.

[Legal Guide to Administrative Data Sharing for Economic and Workforce Development \(State Data Sharing Initiative\)](#)

A report that outlines common legal issues and potential processes to solve them in the context of sharing data for economic and workforce development.

[Data Sharing Resources \(State Data Sharing Initiative\)](#)

A listing of the states that have passed data-sharing legislation and set up data management bodies and MOUs to address legal uncertainties.

Resources (organizational barriers)

[Becoming a Knowledge-Sharing Organization \(World Bank Group\)](#)

A helpful guide to navigating organizational barriers. Begin at page 12 of handbook, at section “Why Are Leadership and a Knowledge-Sharing Culture Important?”

[Building Trust for Cross-Sector Data Collaboration \(National Center for Complex Health and Social Needs\)](#)

A report on the importance of building trust when sharing healthcare data that revolves around six core strategies.

[Access to New Data Sources for Statistics: Business Models and Incentives for the Corporate Sector \(OECD Statistics Working Papers\)](#)

A discussion from Klein and Verhulst of the risks of and incentives for sharing corporate data; begin at page 11.

[What organizations need in order to share more data \(Open Data Institute\)](#)

A summary of insights from interviewing different actors in various sectors about the barriers and opportunities of sharing data.

[Data Donation: Sharing Personal Data for Public Good? \(Digital Economy All Hands Meeting\)](#)

A conference paper by Skatova, Goulding, and ng that presents two case studies of sharing health data for research and finds that organizations that share data for self-benefit act very differently than those who do so prosocially. Public perception of the receiving research organization was also found to be an important factor.

[Incentivizing Data Donation \(Nature Biotechnology\)](#)

An article on the importance of creating incentives to get people to share their health data and some of the current problems with this data-sharing landscape; highlights an organization that has emerged to try to address those problems.

[Team Building \(World Health Organization\)](#)

A guide to building teams and preventing failure. See point 10 on page 11, “Why teams fail?”

[Gaining Executive-Level Buy-In for Data Governance Strategy \(AIM Consulting\)](#)

A brief blog post on how to get executive sponsorship for your data-governance strategies and what to do when you get it.

[Data Collaboratives as “Bazaars”? \(Transforming Government: People, Process and Policy\)](#)

A conceptual framework from Susha, Janssen, and Verhulst for thinking about collaboration.

[The Right Data Governance for Your Organization’s Culture \(Dataversity\)](#)

A discussion of what organizational culture means and why it matters for data governance.

Resources (ethical barriers)

[The Ethics of Data Sharing \(Accenture\)](#)

A white paper that covers some of the arguments, concerns, best practices, models for governance, and a framework for ethical data sharing. Includes a sample risk assessment framework for data sharing and sample timeline.

1.3 Understand Data-Sharing Capacity

Considerations

- *Have you taken stock of your data-sharing capacity?*
- *Have technical barriers to data sharing been identified?*
- *Have data literacy programs or processes been created?*
- *Who will receive the shared data?*
- *In what form will the data be shared?*

Resources

[Technology for Civic Data Integration \(MetroLab Network\)](#)

A report that talks through the technical aspects of managing data sharing within the context of setting up a cross-agency integrated data system (IDS). For a discussion of the technical barriers and capacity for data sharing for social impact, see page 7.

[5 Models for Data Stewardship \(SAS Best Practices\)](#)

A resource to help you think through the different kinds of data stewards, the problems with each, and the barriers to creating a data steward role.

[Online Courses \(School of Data\)](#)

A collection of online courses on data science available to the public.

[We Are the Data Literacy Project \(Data Literacy Project\)](#)

A collaborative effort to help society become more data literate by encouraging leading organizations to provide data literacy training, creating a library of resource-based learning for data literacy, and supporting global *tion should be employed to protect personal data?*

[Data Literacy Program \(Qlik\)](#)

An online resource for data literacy.

What might the recovery process look like?

Resources

[Technology for Civic Data Integration](#) educational institutions to place data literacy into their curricula.

[College & University Data Science Degrees \(DataScience.Community\)](#)

A listing of all data science academic programs in the U.S.

1.4 Understand the Data Being Shared

Considerations

- *What is the quality of the data?*
- *Have appropriate data standards been identified?*

Resources

[Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries \(Microsoft\)](#)

A remarkably complete listing of potential quality issues of social data.

[Five Principles for Applying Data Science for Social Good \(O'Reilly\)](#)

A blog post on best practice principles for using big data for the social good, written by Jake Porway from DataKind. The first principle discusses how the way that organizations have published data might not be all that useful.

[Connecting the Dots: A Data Sharing Framework for the Local Public Health System \(National Association of County & City Health Officials\)](#)

A guide for data sharing for local public health. See page 3, table 2: “Elements of a Data Management Plan.”

[Interoperability and Data Sharing \(OPRE\)](#)

A portfolio that discusses the benefits of collaboration and interoperability

1.5 Survey Security and Privacy Related to the Data

Considerations

- *How sensitive are the data?*
- *How is data encryption being managed through collection, transit, and access?*
- *What are the likely feelings of the relevant communities towards privacy?*
- *What forms of anonymization_*

[Technology for Civic Data Integration \(MetroLab Network\)](#)

This document explores many of the technical aspects of data sharing, with a discussion of security and privacy in this context starting on page 10.

[NNIP’s Resource Guide to Data Governance and Security \(National Neighborhood Indicators Partnership\)](#)

A resource guide on responsible data governance practices; see pages 18–30 for a discussion of good data security practices.

[Data Collaboration for the Common Good \(World Economic Forum\)](#)

See appendix on pages 26–27 for a discussion of centralized versus decentralized (federated) technology solutions to data sharing.

[The “Five Safes”: a framework for planning, designing and evaluating data access solutions \(Zenodo\)](#)

A paper by Felix Ritchie that dives into strategies for ensuring your data access and management approaches allow for security.

[Open Up Corporate Data While Protecting Privacy \(Omidyar Network\)](#)

A blog post by Verhulst and Sangokoya on some of the types of data-sharing partnerships and some of the risks that corporate data-sharing schemes run.

[Nothing to Hide: Tools for Talking \(and Listening\) about Data Privacy for Integrated Data Systems \(Actionable Intelligence for Social Policy\)](#)

A report that talks about data privacy within the context of setting up an intra-governmental IDS system.

[Enabling Data Discoverability, Linkage, and Re-Use \(National Academy of Sciences\)](#)

A workshop summary that discusses how to frame relationships of paramount importance. Rephrasing “data should not be used for research or linked unless it can be done safely and securely” into “data should be available for research and linking unless it cannot be done safely and securely” can significantly increase support.

[Data Protection: Data In Transit vs. Data At Rest \(Digital Guardian\)](#)

Web resource that provides a basic introduction to the different considerations and barriers to securing data while in transit or at rest.

[Minimizing Disclosure Risk In HHS Open Data Initiatives. C. The Mosaic Effect \(U.S. Department of Health & Human Services\)](#)

A discussion of how seemingly unrelated and “non-sensitive” bits of data can be recombined in ways that generate sensitive data to form the mosaic effect.

[Sharing Data for Better Results \(National League of Cities\)](#)

A summary of major federal laws surrounding education, health, social services, and crime data: see pages 16–47.

[Are Data Sharing and Privacy Protection Mutually Exclusive? \(Cell\)](#)

An article by Joly, Dyke, Knoppers, and Pastinen on trade-offs between privacy and data sharing in the context of health data.

[POTs: Productive Optimization Technologies \(Cornell University\)](#)

A study that provides a solution to the privacy concerns of optimization technology through a framework that allows individuals to protect against unwanted consequences.

[How to Use SFTP to Securely Transfer Files with a Remote Server \(Digital Ocean\)](#)

A summary of SFTP along with how to best implement it.

[The Legacy of Inbloom \(Data Society\)](#)

A case study on the story of educational technology in New York through Inbloom

1.6 Examine the Ethical Implications of Sharing the Data

Considerations

- *Who is impacted and what ethical considerations does this population or its data pose?*
- *Has the urgency of the problem been weighed against the potential harm of the solution?*
- *How can the community be engaged from the very early stages?*

Resources

[Data Collaboration for the Common Good \(World Economic Forum\)](#)

Helpful background on public private partnerships. See page 8 for information on the issue.

[Reimagining Measurement: Enhancing Social Impact through Better Monitoring, Evaluation, and Learning \(Deloitte\)](#)

A discussion of how to re-imagine your approach to focus on measures that reflect serving the community of interest and not just your funder’s organizational goals.

[Machine Bias: Investigating Algorithmic Injustice \(ProPublica\)](#)

A collection of stories investigating the ways in which algorithms negatively affect society by perpetuating injustice and bias.

[How Big Data Is “Automating Inequality” \(New York Times\)](#)

Review of a book by Virginia Eubanks that argues that instead of technology increasing the efficiency of government services it actually increases the inequality of said services.

[Weapons of Math Destruction: Cathy O’Neil Adds Up the Damage of Algorithms \(The Guardian\)](#)

An interview where Cathy O’Neil discusses her book about how algorithms can cause more damage to society than we think.

[Civil Rights, Big Data, and Our Algorithmic Future: A September 2014 Report on Social Justice and Technology \(Upturn\)](#)

A report on the social justice and bias issues of modern technology.

[Reforming the U.S. Approach to Data Protection and Privacy \(Council on Foreign Relations\)](#)

A brief overview of why the current approach to regulating privacy and consumer data in the U.S. does not create meaningful protection for individuals.

[Ethics Washing Is When Ethics is a Substitute for Regulation \(DataEthics\)](#)

A discussion of the concept of “ethics washing.”

[Community Advisory Boards \(West Virginia Clinical & Translational Science Institute\)](#)

A brief presentation of what a CAB is and when it makes sense to create and use one.

[The Seductive Power of “Solving” Bias in Artificial Intelligence \(OneZero\)](#)

A discussion of the effects of bias in AI, the best solutions to these problems, and the implications of the data associated with these technologies.

[Why We Need Community-Driven Tech Ethics: Two Years of the Global Data Ethics Project \(Tech at Bloomberg\)](#)

A summary of the growing movement of data ethics and the Global Data Ethics along with the importance of community involvement.

[Oxfam Responsible Data Program Policy \(Oxfam\)](#)

A Phase One exemplar that lists data rights for users.

1.7 Draft your Charter

Considerations

- *What is the purpose of your collective?*
- *What is the context and background of the collective and impacted populations and stakeholders? What are their values and motivations?*
- *Have the most salient legal, security, privacy, organizational, and ethical considerations been acknowledged?*

- *What is the scope of the project and the goals of the collective?*
- *Who are the members of the collective, their roles, and points of expertise?*
- *Is there an authority or a way to confer authority on specific actors to ensure effective operations?*
- *What are the operational plan, decision-making mechanisms, membership change procedures, rule-making processes?*
- *How will performance be assessed?*
- *What timeline and milestones are in place to guide performance assessments and how long will the collective operate?*

Resources

[Creating a Charter for Your Policy Team \(National Institute of Corrections\)](#)

A roadmap that helps policy teams think through and draft a charter.

[Project Charter Template \(Centers for Disease Control and Prevention\)](#)

A template for a team charter.

[Generic Team Charter Template \(Acquire.gov\)](#)

A template for a team charter (warning: direct download from Wayback Machine).

[Acquisition Team Charter Template \(GA Tech\)](#)

A template for a team charter.

[MITA Governance Charter \(Medicaid\)](#)

A template for a team governance charter.

PHASE TWO: DEFINE THE OPERATIONS

2.1 Determine Your Governance Framework Structure

Considerations

- *What models for data sharing support your work?*
- *Which model best fits your needs?*
- *How will the public and affected communities be informed and engaged when determining the data-sharing governance process?*
- *Who will receive the shared data, and in what form?*

Resources

[Mapping the Next Frontier of Open Data: Corporate Data Sharing \(Medium\)](#)

An essay that describes the current trend in corporate data sharing for social impact and introduces GovLab's taxonomy of types of data-sharing agreements.

[The Emergence of Data Collaboratives...in Numbers \(GovLab\)](#)

A blog post that describes the most common sectors for GovLab's taxonomy of data-sharing agreements.

[How to Scale Data Collaboratives? Key Takeaways from the London Data Stewards Camp \(Medium\)](#)

A blog post that describes a meeting of the Data Stewards Network.

[Data Collaboratives Explorer \(GovLab\)](#)

A listing of all data collaboratives GovLab has found, paired with a brief description, a link to the source, and a classification (that is, which type of data collaborative it is in GovLab's taxonomy).

[Data Maturity Framework Questionnaire \(University of Chicago\)](#)

A checklist of qualities for data teams and assets that you can use to evaluate the maturity of your data from the Center for Data Science & Public Policy.

[An Introduction to Data Collaboratives \(GovLab\)](#)

A brief overview of the different ways to organize data collaboratives.

[The Emergence of Data Collaboratives in Numbers\(GovLab\)](#)

A blog post that details examples of each of the kinds of data collaboratives according to GovLab.

[Structuring a Partnership \(William T. Grant Foundation\)](#)

A (still sparse) platform with guiding questions, examples, and resources for structuring your partnership.

[Data Philanthropy: Unlocking the Power of Private Data for Public Good \(Urban Institute\)](#)

A report describing corporate data-sharing for the social good as a new form of philanthropy, offering up different pathways to creating data-sharing arrangements depending on the sensitivity of the data, and describing two use cases.

[AISP Network Site Case Studies \(Actionable Intelligence for Social Policy\)](#)

A collection of case studies of local and state government-created integrated data systems (IDS): when different agencies, sometimes in partnership with private research entities, share information to form a richer database that allows for better service delivery.

2.2 Formalize Responsible Data-Sharing Practices

Considerations

- *Have the roles and responsibilities of each party been established?*
- *Has feedback from stakeholders been solicited and adapted throughout the defining process?*
- *Have rules around modifying the stakeholder mix of the collective been drafted?*
- *How are data used externally beyond the collective?*

Resources

[Developing Data-Sharing Agreement \(William T. Grant Foundation\)](#)

A (still sparse) platform with guiding questions, examples, and resources for developing a data-sharing agreement.

[Governance \(International Association for Impact Assessment\)](#)

A tip-sheet that briefly describes what to consider and what to do when designing good impact assessment governance.

[Recommendations of the Council on Principles for Public Governance of Public-Private Partnerships \(Organisation for Economic Co-operation and Development\)](#)

A report that gives recommendations on how to design a value-focused institutional framework for effective public-private collaboration.

[A Decision Model for Data Sharing \(13th International Conference on Electronic Government\)](#)

A literature review from a conference paper by Eckartz, Hofman, and Veenstra on data-sharing governance practices that includes three brief case studies in the logistics sector and a data-sharing decision model to help actors identify and address barriers to sharing.

[Mastercard Establishes Principles for Data Responsibility \(Mastercard\)](#)

A press release outlining the company's proposed six data responsibilities (security & privacy, transparency & control, accountability, integrity, innovation, and social impact) for the organization and the digital economy.

2.3 Examine the Sustainability of the Operations

Considerations

- *How is the collective funded?*
- *For what period of time will this collective exist?*
- *Have rules around modifying the stakeholder mix been drafted?*

Resources

[Data Collaboration for the Common Good \(World Economic Forum\)](#)

A resource on public private partnerships. See section 5, page 21: “Economic sustainability and scalability.”

[7 Tools to Help Build Sustainable Data Governance \(Earley Information Science\)](#)

A toolkit with guides and resources on how to build a sustainable data governance model.

[Guide to Data Science and Sustainability \(Discover Data Science\)](#)

An extensive guide from DDS to incorporating sustainability into data science.

2.4 Develop an Ethical Framework for Data Sharing

Considerations

- *How will data governance ensure that ethical concerns are considered, heard, and addressed throughout the project?*
- *How are practitioners thinking about data ethics?*
- *How are practitioners acting on principles of data ethics?*

Resources

[Digital Decisions Tool \(Center for Democracy & Technology\)](#)

An interactive tool that asks a series of ethics and fairness-centric questions along the process of building your data-sharing algorithm.

[Building Frameworks, Setting Standards for Ethical Data Use: Our Conversation with Natalie Evans Harris \(Medium\)](#)

An interview on the importance of community-focused data sharing through ethical standards, educational suggestions on sharing data, and comment on the state of data sharing in society.

[Managing Data Ethics: A Process-Based Approach for CDOs \(Deloitte Insights\)](#)

A summary of past and present practices in managing data ethics along with summaries of concepts and skills for CDOs in the field of data ethics.

[Tech Ethics Curricula \(GoogleDocs spreadsheet\)](#)

An open spreadsheet listing most tech ethics courses available in the U.S.

[Tech’s Ethical “Dark Side”: Harvard, Stanford and Others Want to Address It \(New York Times\)](#)

A summary of certain ethical issues within the tech industry and proposed solutions.

[Data Science Ethics \(Coursera\)](#)

A University of Michigan course on Data Ethics that is open to the public.

[Ethics and Law in Data and Analytics \(EdX\)](#)

A course from Microsoft on Data Ethics that is open to the public.

[Data Ethics Course \(National Forum on Education Statistics\)](#)

A course on data ethics that is open to the public.

[Data Ethics—The New Competitive Advantage \(TechCrunch\)](#)

An article that describes how data ethics can be used by companies for competitive advantage in an era where consumers value companies which behave ethically with their data.

[Data Practices Courseware \(Linux Foundation Project\)](#)

A repository of workshops that non-technical people can take to help them understand the basics of good data practices.

[Ethics and Data Science \(O’Reilly\)](#)

A booklet from Patil, Mason, and Loukides that gives a solid introduction to the concept of data ethics and the urgency with which it needs to be applied to all facets of data science.

[MS and E 234 \(Stanford\)](#)

A course that engages with the difficult ethical challenges in the modern practice of data science.

[Oaths, Pledges and Manifestos: a Master List of Ethical Tech Values \(Medium\)](#)

An overview of ethical initiatives in technology.

[Association for Computing Machinery Code of Ethics and Professional Conduct](#)

An annotated code of ethics for computing professionals that outlines general principles, enumerates responsibilities, and offers advice for leadership.

[Ten Simple Rules for Responsible Big Data Sharing](#)

An annotated 10-point framework of key data ethics principles.

[American Statistical Association’s Ethical Guidelines for Statistical Practice](#)

A detailed list of ethical principles aimed at statistical professionals centered around promoting integrity and highlighting responsibilities.

[Data Science Oath of National Academies of Science, Engineering, and Medicine](#)

An actual oath that data practitioners can take, situated alongside the Hippocratic oath.

[Manifesto for Data Practices of data.world](#)

A simple 12-point list of ethical principles centered around four values: inclusion, experimentation, accountability, and impact.

2.5 Define What Success Looks Like

Considerations

- *What impact metrics are appropriate?*
- *How are data used externally beyond the collective?*

Resources

[Reimagining Measurement: Enhancing Social Impact through Better Monitoring, Evaluation, and Learning \(Deloitte\)](#)

A white paper discussing the importance and difficulties in creating good impact evaluations.

[Toward Metrics for \(Re\)Imagining Governance \(GovLab\)](#)

A white paper discussing the use of metrics in evaluating interventions and governance innovations.

[International Principles for Social Impact Assessment \(Impact Assessment and Project Appraisal\)](#)

An article that defines social impact measurement and provides guidelines and values for it.

[Assessing Significance in Impact Assessment of Projects \(International Association for Impact Assessment\)](#)

A tip sheet that defines and provides guidelines for measuring the scale/significance of social impact for specific projects.

[Situating the Next Generation of Impact Measurement and Evaluation for Impact Investing \(IssueLab\)](#)

A white paper promotes a convergence of methods, building from both the impact investment and evaluation fields.

[Policy Brief on Social Impact Measurement for Social Enterprises \(Organisation for Economic Co-operation and Development\)](#)

A white paper that seeks to define exactly what social impact is.

2.6 Draft Data-Sharing Agreement

Considerations

- *What data will be used, how will they be shared and used?*
- *Which stakeholders are involved in what part of the data collection, processing, and analysis?*
- *What security and ethical safeguards have been put in place?*
- *How are collective-wide decisions made and what process do these decisions go through?*
- *When are confidentiality agreements required and how are they drafted?*

- *How are ethical and security reviews and audits conducted?*
- *How are external requests for data or insights including review, approval, and pricing handled?*
- *Who contributes what resources to the project?*
- *Who owns the intellectual property of the data and of the resulting publications and who decides whether to publish or share them externally?*
- *Who gets credit for what part of the collective?*
- *How are potentially ambiguous terms that might interfere with mutual understanding of the contract defined, especially if the partners are from different sectors, been clearly defined?*
- *How are changes that members want addressed?*

Resources

[Developing Data-Sharing Agreement \(William T. Grant Foundation\)](#)

A (still sparse) platform with guiding questions, examples, and resources for developing a data-sharing agreement.

[NNIP Concept \(National Neighborhood Indicators Partnership\)](#)

A repository of example documents, including data-sharing agreements, codification of various security and privacy procedures, and confidentiality agreements.

[Data-Sharing: Creating Agreements in Support of Community-Academic Partnerships \(Colorado Clinical and Translational Sciences Institute\)](#)

A report that outlines best practices, processes, recommendations, and a set of resources to help stakeholders navigate data sharing, data management plans, and data-sharing agreements in the context of community-academic partnerships.

[Guidelines for Developing Data Sharing Agreements to Use State Administrative Data for Early Care and Education Research \(U.S. Department of Health & Human Services\)](#)

A report that walks readers through the creation of data-sharing agreements that seek to repurpose administrative data from state agencies for health-sector research.

[Introducing Contracts for Collaboration: New Project on Legal Conditions for Data Sharing \(Thematic Research Network on Data and Statistics\)](#)

News article that previews an upcoming project that seeks to create an online library of data-sharing agreements accessible to those seeking to draft their own.

[Data Trust Agreement Template \(BrightHive\)](#)

A sample data trust agreement.

[Putting the Trust in Data Trusts \(Register Dynamics\)](#)

A summary of the benefits of data trusts along with sample case studies.

PHASE THREE: DRIVE IMPACT

3.1 Reinforce, Update, and Share Governance Best Practices

Considerations

- *How can you ensure that parties continue to adhere to previously established standards and practices?*
- *How can data-sharing agreements be changed based on lessons you learn from evaluating progress on each step of the theory of change?*
- *How do you decide on and take on new research projects?*
- *How can iterative improvements best be documented and publicized to stay transparent, help other organizations learn from mistakes and failures, and enrich the literature to promote data-sharing for social impact elsewhere?*
- *How do we continue to minimize and assess risk as we continue working?*

Resources

[Iterative Development \(Code for America\)](#)

A discussion of what iterative development means along with a list of tools and resources.

[Rapid Evaluation Approaches for Complex Initiatives \(ASPE\)](#)

This paper builds a framework for how to evaluate different projects based on their levels of complexities.

[Closing the Feedback Loop\(SSIR\)](#)

A program that gives advice on community engagement.

[Effective Stakeholder Engagement \(International Association for Impact Assessment\)](#)

A quick summary of tips and advice on stakeholder engagement from planning to things to know and things to do.

[Public Participation: International Best Practice Principles \(International Association for Impact Assessment\)](#)

A brief guide for how to integrate public participation into your impact assessment including a summary of objectives and best practices.

[Sharing Results: A Guide for Communicating Promise Neighborhoods Outcomes to Diverse Audiences \(Urban Institute\)](#)

A paper that walks you through how to properly frame and convey your results to a variety of audiences.

[Data Walks: An Innovative Way to Share Data with Communities \(Urban Institute\)](#)

A report that helps researchers design ways to share data and findings with community residents and program participants that ensure better interpretation and dissemination of the results and inspire action within the community of interest.

[Evaluating Partnerships for Improvement and Impact \(William T. Grant Foundation\)](#)

A (still sparse) platform with guiding questions, examples, and resources for structuring your partnership.

[Nothing to Hide: Tools for Talking \(and Listening\) About Data Privacy for Integrated Data Systems \(Actionable Intelligence for Social Policy\)](#)

A report that talks about data privacy within the context of setting up an intra-governmental IDS system.

[IDS Case Study: Los Angeles County \(Actionable Intelligence for Social Policy\)](#)

A report on IDS organization in Los Angeles describing how a governing body meets annually to turn research ideas into a research agenda, where agency staff can provide the county board with data sources and information that creates context around the county's initiatives.

[The Role of Case Studies in Effective Data Sharing, Reuse and Impact \(IASSIST Quarterly\)](#)

A discussion about the importance of case studies in encouraging data sharing, noting that there is currently a dearth of well-documented studies in this area that can help practitioners work through this process.

[Cross-Sector Data in Action \(National Center for Complex Health and Social Needs\)](#)

A discussion of a few case studies on the topic of cross-sector data sharing.

3.2 Reinforce, Update, and Share Best Practices in Ethics Governance

Considerations

- *Who are the affected communities?*
- *Are there new externalities and have previously identified ones been properly addressed?*
- *Has technology changed in a way that increases the potential for data misuse?*
- *Can fail-safes be implemented to prevent misuse of data?*
- *Is conducting an equity audit appropriate?*

Resources

[Ethics and Data Science \(O'Reilly\)](#)

A booklet by Loukides, Mason, and Patil on data ethics and how to incorporate these standards into your everyday work life.

[Community Engagement Toolkit \(Collective Impact Forum\)](#)

A tool kit to effectively engage community for data-driven solutions.

3.3 Monitor and Assess Privacy and Security Approaches

Considerations

- *Has the structure of the collective changed?*
- *Has the technological environment changed?*
- *Has the regulatory environment changed?*
- *Has the organizational culture of any stakeholder organization changed?*
- *Has society changed in ways that affect the impacted communities?*
- *Are there system processes that can be put in place to minimize or resolve human errors?*

Resources

[CISO Tips: How CISOs Can Reduce Risk \(CIO\)](#)

An article providing a framework for CISOs of organizations to tackle the evolving technological and privacy environment with five quick tips.

[Advice for CISOs: Want More Resources? Think Beyond Pure Tech \(Symantec\)](#)

A guide for how CISOs can increase security budgets and further tackle privacy and security concerns.

[Considerations for Using Data Responsibly at USAID \(USAID\)](#)

A Phase Three exemplar that shows how to best assess risk.

3.4 Develop Processes to Improve Data Quality Over Time

Considerations

- *What metadata can be collected to make the data more useful?*
- *How can the documentation process be improved and can insight about potential bias be baked in?*
- *How can capacity be built over time, whether technical infrastructure or data management expertise?*
- *How has the situation or your understanding of the situation changed [in a way that changes](#) your idea of effective impact metrics? Can it be changed to improve benchmarking and comparisons with other efforts?*

Resources

[5 Hands-On Strategies to Improve Data Quality \(RingLead\)](#)

A quick tips guide on five strategies for improving data quality over time.

[What is Metadata and Why is it as Important as the Data Itself? \(Opendatasoft\)](#)

This article provides an overview of metadata and how to best use it to our advantage.

[Scoping \(Fast Tips\)](#)

A definition and discussion of what scoping means.



GEORGETOWN UNIVERSITY