

DATA ETHICS TOOLKIT

A Toolkit for Data Ethics in
the Participatory Sciences

Editors

Caren Cooper

Lisa M. Rasmussen

Elizabeth Jones

Acknowledgments

The Editors would like to acknowledge, with our deepest gratitude, the many members of the participatory sciences community who made this project possible. Individuals mentioned below, in addition to approximately fifty individuals who contributed but wish to remain anonymous, have been integral to the project. Thanks to their input, the Toolkit reflects both areas of broad agreement and areas of divergent priorities and values. The Editors have endeavored to avoid, but take responsibility for errors, oversights, and omissions.

Advisors

Anne Bowser
Jean Goodwin
Alison Parker
Toby Schonfeld
Jennifer Shirk
Sacoby Wilson
Michael Zimmer

Consultant-Collaborators

Amy Freitag
Vincent Martin
Omega Wilson

Co-Facilitation Team

Omar Muhammad
Nayyirah Shariff

Community

Engagement Partner

The Citizen Science Association

Support & Design

Reanna Putnam

Original Artwork

Lila Higgins

Trustworthy Data Practices Team (plus 70 who remain anonymous)

| | | |
|---------------------------|---------------------|-----------------|
| Guillermina Actis | Craig Hood | Jack Nunn |
| Lauren Anderson | Barb Horn | Sean O'Connor |
| Maria-Antonietta Buccheri | Pen-Yuan Hsing | Shane Orchard |
| Brandon Budnicki | Michael Jabot | Alison Parker |
| Lydia Collins | Devanshu Jha | Pammi Price |
| Alexandra Czegledi | Xeni Kechagioglou | Turam Purty |
| Veronica Del Bianco | Skye Kelty | Rosemary Race |
| Ursula Ellis | Ankit Khanal | Lea Shanley |
| Jennifer Epstein | Holli Kohl | Nancy Sheehan |
| Justin W. Flory | Jackie Lindsey | Anne Simonis |
| Alexis Garretson | Jaxson Mack | Ted Smith |
| Carlos Goller | Rachael Mady | Rob Stevenson |
| Yaela Golumbic | Lauren Marino Perez | Gefion Thuermer |
| Debbie Gonzalez Canada | Karen Martin | Hossein Vahidi |
| Bastian Greshake Tzovaras | Charmel Menzel | Mariana Varese |
| Muki Halkas | Peder Nelson | Kristen Weaver |
| Barbara Heinisch | Noah Newman | Adam Woods |

Environmental Justice Practitioners Team Members

| | |
|--------------------------|------------------|
| Deborah Stewart Anderson | Erin Proven |
| Chris Canaly | Eloise Thomas |
| Sala Faruq | Lucas K Thornton |
| Christopher Gross | Ayo Wilson |
| Valerie Ann Johnson | Brenda Wilson |
| Gloria J Lowe | Omega Wilson |
| Vincent Martin | |



This material is based upon work supported by the National Science Foundation under Grant No. [1835352], “Establishing Norms of Data Ethics in Citizen Science.”

HOW TO CITE THIS DOCUMENT: Cooper, C. B., Rasmussen, L. M., and Jones, E. D. 2022. A Toolkit for Data Ethics in the Participatory Sciences. Citizen Science Association.

TABLE OF CONTENTS

- 3 ACKNOWLEDGMENTS**
- 5 INTRODUCTION**
 - 1. Welcome
 - 2. Ethics Primer
 - 3. Toolkit Structure and Outline
 - 4. Concluding Thoughts
- 11 DATA GOVERNANCE: KEY CONCEPTS**
 - 1. Participatory Data Streams
 - 2. Ethical Frameworks for Data Governance
 - 3. Positionality within Data Governance
 - 4. Governance Structures
- 17 DATA GOVERNANCE: WORKSHEET FOR YOUR PROJECT**
- 21 DATA INTEGRITY: KEY CONCEPTS**
 - 1. Scientific Integrity and Misconduct
 - 2. Data Quality in Context
 - 3. Conflicts of Interest and Commitment
- 23 DATA INTEGRITY: WORKSHEET FOR YOUR PROJECT**
- 25 REPORT-OUTS: KEY CONCEPTS**
 - 1. Benefits of Report-Outs
 - 2. Risks & Tensions with Report-Outs
 - 3. Preparing Report-Outs
- 29 REPORT-OUTS: WORKSHEET FOR YOUR PROJECT**
- 31 RECOMPENSE: KEY CONCEPTS**
 - 1. THE ROLE OF RECOMPENSE
 - 2. APPROACHES TO RECOMPENSE
- 35 RECOMPENSE: WORKSHEET FOR YOUR PROJECT**
- 37 REPORT-BACKS: KEY CONCEPTS**
 - 1. BENEFITS OF REPORT-BACKS
 - 2. REPORT-BACK FORMATS
- 39 REPORT-BACKS: WORKSHEET FOR YOUR PROJECT**
- 41 TRUSTWORTHY DATA PRACTICES SUMMARY**

1) WELCOME

Welcome to the Data Ethics in the Participatory Sciences Toolkit, a resource developed collaboratively by and for practitioners of citizen and community science. While many aspects of participatory science involve ethical considerations, this toolkit is built for considering ethical issues surrounding data, the bedrock of science. This introduction describes the ethical foundations of the toolkit; its scope and genesis; and brief outlines of the tools. While significant time, attention, and other resources have been invested in the toolkit, we view it as a living document and hope the community of practice will continue to develop it and add additional tools.

A. Project Genesis and Process

This project grew over the course of several years. In 2015, at the first annual conference of the Citizen Science Association (CSA), Cooper led a symposium on ethics and subsequently launched the Ethics Working Group. These two actions planted the seeds for the Toolkit. The idea began to sprout from two other efforts: Rasmussen held a mini-conference on ethics in citizen science in 2017, and Cooper co-chaired an international group focused on data curation of volunteer-generated data (2017–2020). In 2019, Cooper and Rasmussen were awarded an NSF grant to support the co-creation of the Toolkit, with the effort coordinated by Post-Doctoral researcher Elizabeth Jones.

We designed the Toolkit to support ethical thinking, but there are inherent limits to the efficacy of any toolkit or training about ethics. A culture of ethical norms, promoted by a peer community of practice, is a key component to bring about ethical behaviors. Therefore, we designed the Toolkit as a collaborative effort with the Citizen Science Association, a community of practice spanning many forms of participatory science. Our intention is to support the CSA's capacity to sustain a culture of data ethics in the participatory sciences.

The Trustworthy Data Practices Team consisted of over 130 individuals from over a dozen countries and with expertise across a dozen sectors and multiple scientific disciplines who gave input into the creation of this toolkit. Participation involved several CSA working groups (WGs), including the Environmental Justice Practitioners WG, the Data & Metadata WG, the Ethics WG, the Law & Policy WG, and the Integrity, Diversity & Equity WG. As the Trustworthy Data Practices Team, these individuals participated in surveys, focus groups, and workshops over the course of two years to help explore and identify norms of data ethics in the participatory sciences. An initial series of focus groups and workshops lacked representation of grassroots environmental justice practitioners. Therefore, we organized a second series that engaged, and was co-facilitated by, environmental justice practitioners. Further, we hope that the ongoing discussion and growth of the Toolkit by CSA working groups, affinity groups, and members, will sustain the norms of trustworthy data practices.

See citizenscience.org/data-ethics for additional details on the history and process of this project.

B. Toolkit Audience and Scope

Participatory science projects come in nearly limitless forms, but can share some key features. This Toolkit mainly provides guidance for project leaders of so-called “top-down” projects, driven by institutions, with the common feature of centralized data management and stewardship. The Toolkit was co-created with input from these practitioners as well as from individuals with expertise and experience with bottom-up, community-driven projects. Because project leaders bear the responsibility for centralized data stewardship and the execution of data decisions (even when they do not make those decisions by themselves), the Toolkit also focuses on their role in identifying and satisfying, or sometimes balancing, ethical obligations

of the project to participants, partners, science, and society. The Toolkit considers obligations among three categories of people described below.

Roles relevant to considerations of ethical obligations

Project Leaders design, implement, and oversee a research project. They execute project decisions, and whenever possible and appropriate, involve participants and/or partners in those decisions.

Participants contribute data to a project, and may engage in a variety of other ways, from deciding on a research question to contributing, analyzing, interpreting, or visualizing data. They are a heterogenous group, varying by motivations, interests/agendas, or social identities, among other things.

Partners do not necessarily contribute data, but may collaborate in research, help project leaders with decisions about project design, objectives, and activities, and/or connect leaders with interested participants. Like participants, partners are a heterogenous group who may appropriately play a role in project decisions, depending on the nature of collaboration.

The Toolkit content is mainly US-focused, although the development of the Toolkit involved members of the international community of practice. Keep in mind that your project may also be subject to laws, rules, and guidelines not covered in this Toolkit. For example, depending on the project activities and where they take place, human subject regulations, animal care regulations, biosafety guidelines, HIPAA, COPPA/CIPA, GDPR, or other national or international laws may apply.

The tools are designed not to prescribe specific or singular answers, but instead to support reflection and iteration on project decisions. The tools can therefore support the establishment of new projects and the transformation of existing projects. We encourage you to use these tools to help you think like an ethicist within the specific context of your own project.

The Toolkit focuses on data ethics, which therefore leaves out many other types of ethical issues that can arise from participatory science projects.

C. Thinking Like an Ethicist

The Toolkit focuses on data ethics because all participatory science projects depend on data shared or generated by participants and/or partners. These data are assembled by many and held centrally, which leads to unique obligations for stewarding and disseminating data in a way that fosters trust in the participatory sciences. Ensuring trustworthy data is not based solely on technical decisions; it depends on ethical ones too.

The Data Ethics in the Participatory Sciences Toolkit is a primer to help project leaders like you consider ethical obligations regarding data generated by participatory science projects. Ethical explorations are different than scientific ones. A scientific question can result in an agreed-upon scientific answer. In contrast, an ethical question can result in many reasonable ethical answers. What arises as an ethical issue and appropriate solution in one project might not in another, almost identical project. The toolkit guides project leaders to 'think like an ethicist' by helping them to identify ethical issues and obligations, and become familiar with how to use key ethical concepts and principles to make appropriate decisions for their project.

2) ETHICS PRIMER

Because the Toolkit is not designed to result in a single ethical answer for all projects, learning about some general ethical principles will be helpful as you “think like an ethicist.” Ethical principles are a way of reminding ourselves about our basic values and commitments as we consider a situation’s ethical dimensions, and can be used as lenses directing attention to moral features of a situation. So, although ethical principles can’t be used to calculate a specific answer to a moral question, they focus thoughts on such aims as identifying unacceptable solutions, not overlooking important considerations, or clarifying reasons for disagreement between decision makers. Below are some of the ethical principles that play important roles in medical and behavioral research, data science, and other areas relevant to the participatory sciences, as well as examples of the obligations they convey.

Respect is the principle of treating individuals as worthy of moral consideration and as the appropriate decision makers about what happens to them. In medicine, this means that physicians and researchers have an obligation to ask patients for their consent in treatment and research. In the participatory sciences, respect means project leaders have an obligation to learn about and incorporate the influence of participant preferences and priorities regarding data practices and project outcomes.

Reciprocity is a principle of fairness in exchanges of mutual advantage. This principle can help avoid exploitation and maintain a “two-way street” in cases where power inequities are present. In the participatory sciences, it means a duty to structure projects so that participants accrue benefits, instead of taking their efforts for granted.

Transparency means providing clarity and cultivating openness and understanding regarding important information. For example, in reporting research results, transparency requires disclosure of potential conflicts of interest. In the participatory sciences, transparency implies that project leaders have a duty to communicate intentionally with participants and partners about project decisions and their implications regarding how data are managed, shared, used, and potentially re-used.

Accountability means accepting responsibility for both one’s work and one’s conduct with others. In research, authorship requires taking accountability for the reported results. In the participatory sciences, project leaders are responsible for things such as minimizing harms, maximizing benefits, ensuring the research integrity of themselves and participants, and achieving stated project objectives.

You can consider additional obligations derived from these principles as well as find additional principles important to your work. Once you have articulated a set of guiding moral principles and their corresponding obligations, you may find that there are tensions between obligations. It is often impossible to satisfy all obligations simultaneously, so identifying one’s duty in a particular situation may require evaluating which obligations take priority or how to balance them, whether obligations arise from within a single principle or between principles.

In the participatory sciences, tensions between obligations may arise because of the unique fact that participants generate the data. They contribute data with personal details, the research outcomes may affect their lives, and they contribute these data to a central place where others execute decisions about that data. Consequently, participants must place trust in others to treat the data responsibly and ethically.

As a consequence of this prevalent tension in the participatory sciences, the structure of this toolkit focuses on two main sets of obligations: to the participants who make it possible, and to science. These two obligations can generate tensions if obligations to science and participants cannot both be satisfied simultaneously, and projects may strike the balance differently.

Obligations to project participants: Leaders of a project have a responsibility to protect and steward data for the benefit of the participants who contribute it and make the research possible (See Recompense Tool), as well as a responsibility to share back the results in forms they can use (See Report-Back Tool).

Obligations to science: The participatory sciences are sciences, so they share in the obligations to conduct research with rigor and integrity (see Data Integrity Tool), and to share the results of research with the scientific community (see Report-Out Tool).

When initial plans for managing and sharing data in science stem from what is best for the research team, without consideration of other obligations, the result can be harmful, extractive practice. Instead, in their role as executing decisions about data on behalf of all who made the dataset possible, project leaders can select governance structures that broaden decision-making (see Data Governance Tool).

3) TOOLKIT STRUCTURE AND OUTLINE

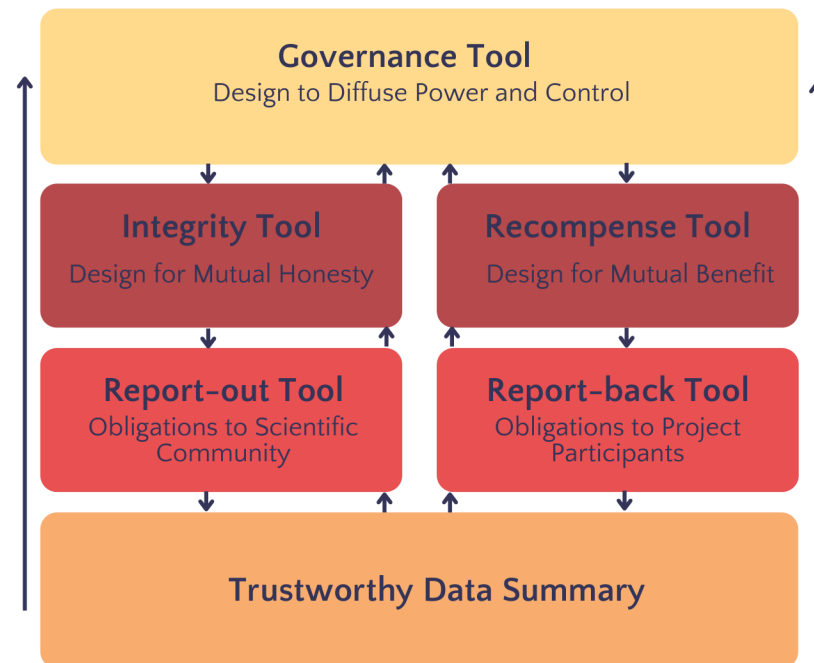
The remainder of this introduction briefly outlines the organization of the Toolkit.

Each tool includes the following paired items:

- **Key Concepts:** defines the key ideas, explains why they matter, offers useful frameworks or concepts to help identify obligations and approach decisions, and concludes with a bottom line. Occasionally you will be prompted to ‘pause and think’ about examples relevant to the field.
- **Worksheet:** a series of prompts to help explore ethical tensions and identify the solutions that are appropriate and achievable for a given project. Each worksheet also prompts iteration to increasingly broaden considerations of ethical obligations.

Whether you are using the toolkit while designing a new project or after a project has begun, start with the first tool, “Data Governance,” because each supporting tool draws on this initial foundational tool as outlined below.

Toolkit organization and interrelationships



Foundational Tool: Data Governance

The Toolkit begins with a fundamental tool, Data Governance, which helps project leaders position their role as executing decisions about data on behalf of all who made the dataset possible. Through prompts to help identify who, how, and why individuals could be affected by data governance decisions, the tool also prompts options for who, how, and why their interests can influence decisions. This tool and its accompanying worksheet prompts you with questions to consider as you build your data governance plan, and helps you consider the obligation to communicate those decisions transparently with project participants and partners. The worksheet in each subsequent Supporting tool draws on the Data Governance tool, allowing for continued iterations.

Supporting Tools

A number of mutually intertwined supporting tools explore more specific areas of data ethics. For example, the “Data Integrity” tool functions as a precursor to the “Report-Outs” tool, and the “Recompense” tool functions as a precursor to the “Report-Backs” tool. The worksheets in the Supporting Tools draw on Data Governance, allowing for continued iterations.

- The Data Integrity Tool will help project leaders with the complexities of data integrity given that participants, partners, and project leaders share in this ethical obligation.
- The Report-Outs Tool addresses the description and sharing of data and its documentation with the scientific community for use and re-use.
- The Recompense Tool aims at identifying the best way to make participants’ efforts on a project worthwhile to them. This can range from financial compensation to professional development, authorship on publications, solutions to environmental concerns, or a combination.
- The Report-Backs Tool focuses on sharing project data back to participants and partners in ways that are understandable and usable to them.

Final Output Tool

When used holistically, the final output of the toolkit is a Trustworthy Data Practices Summary, where project leaders can record decisions made using the tools. The content of the Summary can also serve as the basis for communicating such decisions to participants and partners, or more broadly, as part of the obligation to be intentional and transparent about your project’s data practices.

4) CONCLUDING THOUGHTS

It is never too late to start

No matter where you are in your project journey, the Participatory Sciences Data Ethics Toolkit helps project leaders, along with their participants and/or partners whenever possible, consider or revisit data decisions in their project. Whether starting from scratch or reevaluating a current project, we highly recommend starting with the Data Governance Tool, because deliberation about your decision making structures will affect all other areas of your project.

Next Steps

The Toolkit is a resource that the Citizen Science Association and related communities of practice can leverage to cultivate norms of ethical practices for the participatory sciences. As a co-created set of tools, we thank all who gave their time to the project, and we view this as a living document that individuals, working groups, or affinity groups might refine and expand. You can help establish ethical norms by sharing your experiences with these tools on CSA Connect.

› Data Governance: Key Concepts

What Is Data Governance?

Data governance refers to decision making across the data life cycle to ensure data integrity, availability, usability, and security. All scientific activities that generate data require governance, which includes both technical and ethical dimensions. This tool focuses on the ethical aspects of data governance in the participatory sciences. The ability of the participatory sciences to co-produce new knowledge is grounded in the collection, centralization, and subsequent governance of data.



Data Governance refers to decision making across the data life cycle

Why Does Data Governance Matter?

Project leaders are accountable for decisions about data because of the centralized nature of data collection. To prepare project leaders for a broad array of governance decisions, this tool reviews key concepts in three areas: 1) types of data streams and their potential for harm; 2) ethical frameworks for data governance; and 3) data governance structures.

1. PARTICIPATORY DATA STREAMS

Thinking about participatory science data as falling into three streams will help project leaders identify and consider the ethical obligations and solutions in each. Primary data are the main interest of project leaders, partners, and participants. Enrollment/Administrative data are ancillary to project goals but necessary for managing project activities. Incidental data arise unintentionally from primary data. Each of these may include sensitive data, which are data that could result in harm to people and/or the environment if accessed inappropriately.

Types of Data Streams

Examples:

Primary Data

Data Generated by Participants for Project Goals

- Observations
- Classifications
- Measurements
- Geolocations
- Photographs
- Videos
- Protected Sites
- Protected Species
- Health Records

Enrollment/ Administrative Data

Personally Identifiable Information Provided by Participants for Project Management

- Name/Username
- Mailing Address
- Email Address
- IP Address
- Sex/Gender
- Race/Ethnicity
- Age/Birthdate
- Geolocations
- Photographs

Incidental Data

Personally Identifiable Information Derived from Primary or Enrollment/Administrative Data

- Time-stamped series of geolocated insect observations incidentally reveal participant locations and movements
- Geolocation of lead service lines incidentally reveals households with risk of lead in water
- Time records of annotating videos incidentally reveal when a participant was online

2. ETHICAL FRAMEWORKS FOR DATA GOVERNANCE

Project leaders act as data stewards executing data governance decisions. Making decisions for the data within one's control requires acknowledging the ethical commitments and interests guiding your project. There are several well-developed frameworks that could be applied to the participatory science context. These frameworks embody different, but not necessarily mutually exclusive, ethical commitments, so more than one of these frameworks, or none at all, may be appropriate for your project. Additionally, if your project is hosted through a platform, such as SciStarter, CitSci.org, Anecdata, Zooniverse, iNaturalist, etc., there may be ethical frameworks already adopted and decisions made for the collection, management, and stewardship of project data on that platform. Below are several examples of common ethical frameworks in science.



Open Science

Conventional science is increasingly adopting expectations, incentives, or even requirements to make data open for purposes of re-use. This requires a Data Stewardship mindset towards the long-term care of data as a public good, including ensuring that it is discoverable and usable beyond the immediate interests of anyone associated with the project. In participatory sciences, there may be ethical reasons not to manage data as open, such as if unrestricted re-use of the data might be counter to the interests of participants. Data are not open by default, but require specific steps such as designating data as a public domain resource freely available to everyone, or licensing data for re-use with some restrictions (see Report-Out Tools).



Ownership

Considering data as being owned focuses predominantly on controlling data, which can include both for-profit and not-for-profit purposes. One version of data ownership is to treat data as proprietary, with restrictions on sharing and usage, which can be used by both for-profit and nonprofit organizations. Another approach to data governance is the Community Owned and Managed Research (COMR) model, which prioritizes community ownership and control of data over academic control.



Legal Privacy Restrictions

Although laws and regulations are not the same as ethical frameworks, the former often represent broad public sentiment about the latter. For example, the European Union's General Data Protection Regulation articulates important obligations that constrain making all data open without restriction. If your project might have any participants in European countries, then your project governance must comply with GDPR.

“

Making decisions for the data within one's control requires acknowledging the ethical commitments and interests guiding your project.

”

Addressing Privacy

There are several techniques that can help decrease or remove privacy risks that could harm people and/or the environment. These techniques vary, but include minimization (collecting only the data absolutely necessary for the conduct of the research), obfuscation (blurring of collected data to prevent privacy risks), and security (ensuring that access to information is controlled)



Findable



Accessible



Interoperable



Reusable

FAIR Principles

The FAIR Principles refer to data that are “Findable, Accessible, Interoperable, and Reusable.” These are intended to promote open science and wide sharing of data through standardization of data sets. They may also be useful for purposes other than open science.



Collective Benefit



Authority to Control



Responsibility



Ethics

CARE Principles for Indigenous Data Governance

The CARE Principles are “Collective Benefit, Authority to Control, Responsibility, and Ethics.” These principles arose in response to the FAIR principles “does not fully engage with Indigenous Peoples rights and interests” (<https://www.gida-global.org/care>). For example, in some cases, the need for increased control, especially to enable community benefit or limit how the data are used, is balanced with open science.

3. POSITIONALITY WITHIN DATA GOVERNANCE

Science is often perceived as objective or neutral, but it is actually built upon unstated values. These values are not identical from person to person, are influenced by social identity and lived experiences, and may yield different motives and expectations. Data governance designs should therefore consider how the experiences of those in governance roles may affect their understanding of the ethical responsibilities in a project.

“Positionality” refers to this consideration of how one’s “position,” including social identity relative to a particular context, may limit the breadth of one’s perceptions. Social identity includes, but is not limited to, attributes including gender, sexuality, race, social class, as well as the intersections of those features. It can also include one’s assumptions and beliefs about ways of knowing and discovering, or what constitutes worthwhile research.

In its original form, reflection on one’s positionality typically results in a disclosure statement in a publication for the authors to be transparent about their potential biases. For data governance in the participatory sciences, positionality practices raise project leaders’ awareness about governance structures and power disparities and other ethical tensions. The positionality steps engage project leaders in: 1) reflecting on their lived experiences, social identities, and beliefs and assumptions about knowledge production; and 2) recognizing subjective influences on research given how their values ‘position’ them relative to participants, partners, science, the research topic, and other aspects of the project. The accompanying worksheet invites project leaders to develop their sense of positionality and use it as a lens to reexamine data governance.



Pause and Think

Can you think of ways in which a project leader’s positionality might differ from the positionality of participants and partners?

Tips:

- *The project leader might come from a high-biodiversity area, while participants’ environment may lack biodiversity.*
- *The project leader may be able-bodied while participants may have impairments.*
- *The project leader may prioritize insights from quantitative data while participants value insights from stories and other qualitative data.*

Components of Positionality

Examples:

Social Identity

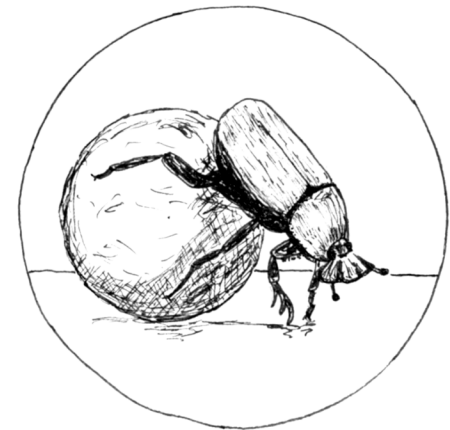
- Class
- Nationality
- Ability
- Race/ Ethnicity
- Age
- Religion
- Sexuality

Lived Experiences

- Education
- Career
- Family
- Wealth/ Income
- Geography
- Social Network
- Institutional Norms

Belief Systems

- Epistemological
- Ontological
- Spiritual
- Political
- Economic
- Moral



4. GOVERNANCE STRUCTURES

Data governance structures specify the roles individuals have in project decision making, and may vary from project to project. Because data collected by participants are centralized by project leaders, those who have collected the data may lose some control of how data are used, resulting in potential disparities in power and responsibility. These disparities can be balanced by structuring a project to diffuse the concentration of decision making power. Designing a data governance structure will involve considering the implications of how power and control of data are distributed among project leaders, participants, and partners.

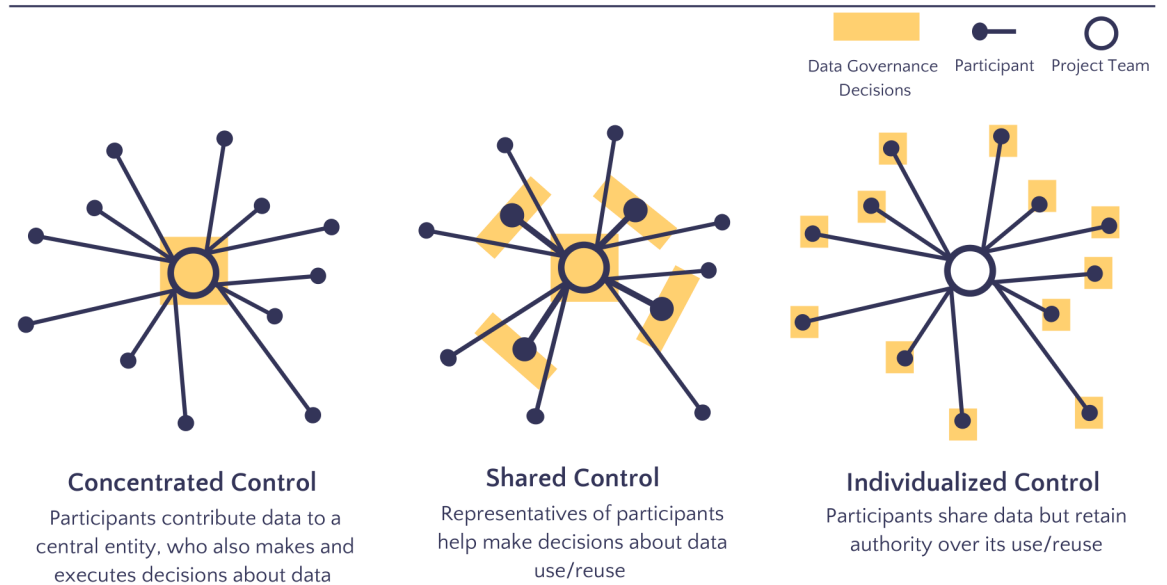


Pause and Think

How can projects be structured so that participants retain some control of their data?

The figure below represents how decision-making power may be shared or limited, and the extent to which decisions are flexible or uniform across participants and/or partners. The simplest structure is where project leaders make and execute uniform decisions for an entire data set. More complex structures include democratic project decision making (shared governance) and greater participant control over their contributed data (individualized control).

Data Governance Decision Making Structures



In the worksheet accompanying this tool, project leaders are invited to decide what general data governance structure is appropriate for their project. Projects may require different governance structures at different decision points, so supporting tools in this toolkit will revisit and build on this foundation.

Bottom Line

Project leaders are data stewards accountable for executing data governance decisions. To identify and address ethical issues in data governance, project leaders should establish governance structures to diffuse the concentration of power and control.

➤ Data Governance: Worksheet For Your Project

PART 1: ESTABLISH FOUNDATIONS

A. Data Streams

1. What are the types of data streams in your project?

Ethical obligations often stem from the types of data collected. In your project, identify which are primary data, which are administrative data, and which are incidental data.

2. Identify any data streams that might be sensitive. If so, would any of the following solutions address the problem?

- Minimize the data collected to only what is necessary for the project
- Obfuscate collected data if they will be shared more broadly
- Ensure data are secure and accessible only to appropriate users based on your data governance plan

B. Ethical Frameworks

1. Have you adopted (or do you aspire to adopt) any of the common ethical frameworks used in science?

Some common ethical frameworks that can be useful to the participatory sciences include the CARE principles, FAIR principles, and open science.

2. Does the framework you've adopted require modification for use in a participatory context?

Consider ways to modify the framework to take into account participant and partner interests, such as in privacy protections and/or openness.

3. What implications do your ethical frameworks have for your data governance plans?

A commitment to open science might require ensuring that your data is findable online, and a commitment to the CARE principles might mean that the data should be used solely for the benefit of the community involved.

C. Consider Potential Constraints

1. Do members of the project team possess the technical skills to achieve the solutions identified?

Skills in geospatial analytics may be necessary to properly obfuscate geolocation data while still presenting data on maps.

2. Are there any legal considerations your project should take into account?

If your project accepts data from countries that belong to the European Union, then data handling must comply with the General Data Protection Regulation (GDPR) law.

3. Are there additional ethical implications of the data collected?

Data collected about contaminated environments, or by marginalized or indigenous communities, may be sensitive.

PART 2. REFLECT ON POSITIONALITY

Research is typically conducted as though there is an objective perspective on the research process. However, a significant body of evidence demonstrates that research and researchers are not value-free. When project leaders reflect on how their social identities and lived experiences relate to the research design, partners, and participants, they can mitigate biases to ensure trustworthy data practices.

By following the steps below, project leaders will reflect on their positionality, which will prime them to identify ethical issues that they might otherwise overlook.

A. DESCRIBE YOUR POSITION

Describe components of positionality, such as social identity, relevant lived experiences, and belief systems:

Components of Positionality

| | Examples: | |
|--------------------------|---|--|
| Social Identity | <ul style="list-style-type: none">• Class• Nationality• Ability• Race/ Ethnicity | <ul style="list-style-type: none">• Age• Religion• Sexuality |
| Lived Experiences | <ul style="list-style-type: none">• Education• Career• Family• Wealth/ Income | <ul style="list-style-type: none">• Geography• Social Network• Institutional Norms |
| Belief Systems | <ul style="list-style-type: none">• Epistemological• Ontological• Spiritual | <ul style="list-style-type: none">• Political• Economic• Moral |

B. Reflect on Your Position

1. How does your positionality relate to the topic of your project?

If your project is about biodiversity, consider whether you live among low or high biodiversity.

2. In what ways are you similar to (an insider) and different from (an outsider) participants?

If your participants tend to be parents, consider how their experiences are similar or different than yours.

3. How might your beliefs and assumptions about knowledge production differ from participants?

As a STEM professional, you may hold the belief that hypothesis-driven, basic research is best for society, while your participants may believe that science can be extractive.

4. How does your positionality inform the research question and other aspects of project design?

Your discipline may prioritize a particular "hot" topic that dictates the research questions.

5. What unstated values and priorities do you hold?

You may view gardening as recreation, and participants may view it as a necessary supplemental food source.

PART 3: SHAPING DATA GOVERNANCE DECISION STRUCTURES

The steps below help you to make decisions about the overall approach to data governance. Other tools will prompt you to consider how to modify the overall approach for particular areas of your project.

1. Given your data governance plan, who is affected (or potentially affected; positively or negatively) by data governance decisions?

2. Considering your positionality, which data governance structure is a good option to enable those affected to influence decisions: scenario 1, 2, or 3 below? In what ways can participants retain some control of their data?



**Scenario 1
(Concentrated control)**

is a frequent default structure where authority and responsibility for governance decisions remain with the data stewards.

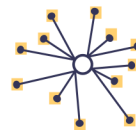
For example, if a project launches without first identifying partners or potential participant groups, then decisions will lack broader input (without shared control) and only have technical capacity to handle participant data uniformly (without options for individualized control).



**Scenario 2
(Shared control)**

allows, in most circumstances, for representative members of different segments of participants to share in decision-making.

For example, cultivating (and compensating) a participant advisory board is a way to incorporate broader input.



**Scenario 3
(Individualized control)**

makes it possible in some circumstances for participants to retain control of the use(s) of their individual contributions, such as permitting them to be used for some, but not other, project purposes.

For example, participants in Patients Like Me can opt to share their data with other participants, but not pharmaceutical researchers.



BRING IT TOGETHER

Based on the work you were invited to do above, including iterating and revising as necessary, summarize your data governance plan here, and include in Box 1 of the Trustworthy Data Practices Summary.

Your positionality and governance structure(s) play a central role in considering ethical obligations in your project. You will return to these considerations in subsequent tools. It is good practice to communicate the project decisions about data governance to your participants and partners.

› Data Integrity: Key Concepts

What is Data Integrity

The concept of data integrity emphasizes the production of accurate, minimally biased data fit for their intended use(s). In the participatory sciences, data integrity requires a shared commitment from project leaders, participants, and partners to honest research practices, and project leaders are responsible for structuring projects to foster these commitments.



Data integrity is the foundation for establishing and maintaining trust in science

Why does Data Integrity Matter?

Confidence in scientific claims arises from trusting that project results are honest and accurate. Professional science establishes this trust through training and participation in scientific institutions and processes. Data integrity deserves particular attention in the participatory sciences because the field is in the process of developing its own practices and institutions to foster trust. Project leaders are responsible for structuring projects to support a climate of data integrity among their project team and participants, such as by demonstrating their own commitments and supporting participants in their practices.

1. SCIENTIFIC INTEGRITY AND MISCONDUCT

Integrity in science refers to honest research practices. Misconduct is a specific type of lack of integrity, including acts of fabrication of data, falsification of data, and plagiarism.

Fabrication

is creating data that does not exist in real-world experiences and investigations.

Falsification

is manipulating research equipment or changing or omitting data to produce inaccurate results (excluding honest mistakes).

Plagiarism

is using another person's ideas or words without giving appropriate credit.



Pause and Think

How might project leaders foster participant commitment to data integrity?

Tips:

- *Demonstrate their own commitment to data integrity.*
- *Design projects to minimize participant conflicts of commitment.*
- *Train participants in these principles, practices, and skills.*

2. DATA QUALITY IN CONTEXT

Data quality is contextual, which means the level of quality depends entirely on the intended use(s) of those data. Quality refers to both accuracy (how closely data reflects truth) and precision (how closely clustered data points are).

Scenario A

Low accuracy | Low precision



might be suitable for conducting pilot studies or generating data while practicing sampling techniques.

Scenario B

Low accuracy | High precision



might be suitable for identifying systematic equipment errors.

Scenario C

High accuracy | Low precision



might be suitable for identifying a particular environmental risk with low-precision instruments.

Scenario D

High accuracy | High precision



would be suitable for most purposes, but some purposes may not require such high precision, particularly if that is more costly.

3. CONFLICTS OF INTEREST AND COMMITMENT

Conflicts of interest or commitment arise when researchers experience two or more incompatible aims. For example, if they may be financially or personally affected by the results of their work, that may introduce pressure to reach certain conclusions.

Examples of conflict of interest or commitment compromising data integrity:

Project Leaders

Pressures to earn consulting contracts may incentivize ignoring sources of bias in the data.

Partners

Desires for protection of a species may incentivize falsification of data.

Participants

Students graded on the number of contributions may be incentivized to fabricate data.



Pause and Think

How might motives and incentives affect commitments to data integrity?

Tips:

- *If participants are competing in a gamified project, might they be prone to dishonest research practices to get ahead?*
- *If participants are recognized for data accuracy, or reaching agreement with other participants through collaborative data-gathering, would that help to encourage data integrity?*

Bottom Line

Trustworthy data are a cornerstone of participatory sciences and dependent on integrity in data generation. Project leaders and participants share the responsibility for data integrity, and leaders are also obligated to ensure that projects are structured to foster and support data integrity.

➤ Data Integrity: Worksheet For Your Project

DEVELOP AND REFINE YOUR DATA INTEGRITY PLAN

Planning for data integrity is often an iterative process. This tool is set up to help you iterate on a data integrity plan by considering important elements, and then considering possible gaps. Though data integrity is a responsibility shared by project leaders and participants, it is the obligation of project leaders to anticipate and plan for how best to support data integrity in the project.

A. Establish Foundations

1. To achieve your intended project goals, what are the key elements of your data integrity plan?

Plan to ensure the appropriate level of data accuracy, minimize bias, ensure equipment calibration, etc.

2. Which parts of the plan above are project leaders directly responsible for, and which are under the purview of participants/contributors?

Project leaders may ensure data are stored securely and confidential, and participants may be primarily responsible for collecting accurate data points.

3. How will you support and communicate the division of responsibility for data integrity and management to participants?

Evaluate the need to provide participants with a data integrity plan and information and/or training on the data collection process.

B. Consider Potential Constraints

1. What practical training is necessary to support the needed level of accuracy and precision?

Evaluate participant data skills and offer training when necessary.

2. What training do participants need on the concepts and norms of scientific integrity?

Evaluate participant understanding of and commitment to data integrity, and offer training when necessary.

3. Are there ways in which project structure and/or participant motivations or conflicts of interest could influence data integrity?

By recognizing these, project leaders can consider how to minimize their potential effects on data integrity.

4. Are there other potential factors that may affect data integrity that need to be addressed?

Structure projects so that resources (such as budget, time, or expertise) permit the data quality required for project goals.



BRING IT TOGETHER

Based on the work you were invited to do above, including iterating and revising as necessary, summarize your project's data integrity plan in Box 2 of the Trustworthy Data Practices Summary. Specify both how the team is assuring data integrity and what the project will provide to the participants to support data integrity.

Report Outs: Key Concepts

What is A Report-Out?

Report-outs are products (e.g., websites, databases, publications, etc.) that share data (typically primary data, not administrative data) and its documentation with the scientific community and those interested in its use and re-use. (This is distinct from report-backs, in which project leaders share data back with participants and partners to help them make meaning of their data contributions.) Report-outs take the form of primary data shared in repositories or archives, often associated with traditional publications that report results and interpretation of the data.



Report-outs satisfy obligations of reciprocity and transparency to the scientific community

Why Do Report-Outs Matter?

Discovering, validating, and understanding scientific knowledge requires sharing and communicating data and results. Science builds on other science, and because participatory sciences are science, they share that obligation. Sharing primary data with the scientific community achieves several goals: it contributes back to the scientific enterprise; honors a commitment to transparency; and conveys important benefits to science, society, and the project itself.

1. BENEFITS OF REPORT-OUTS

Benefits of report-backs flow to science, society, and individual projects.

Scientific Benefits

Sharing primary data can improve scientific research productivity by creating opportunities to reuse, build on, and replicate data.

Project Benefits

Sharing primary data can increase scientific credibility and public trust in the project.

Societal Benefits

Sharing primary data can create greater public awareness and understanding of science, and foster social trust, support, and engagement.



Pause and Think

What are some examples of tensions between obligations to open science and obligations to individual or community participants?

Tips:

- Shared data may reveal locations participants frequently visit, resulting in a loss of participant privacy with feelings of vulnerability to stalking.
- Shared data may create undue financial and reputational burdens due to disclosures of nearby environmental degradation.
- Shared data may be re-used for commercial interests in ways counter to project goals.

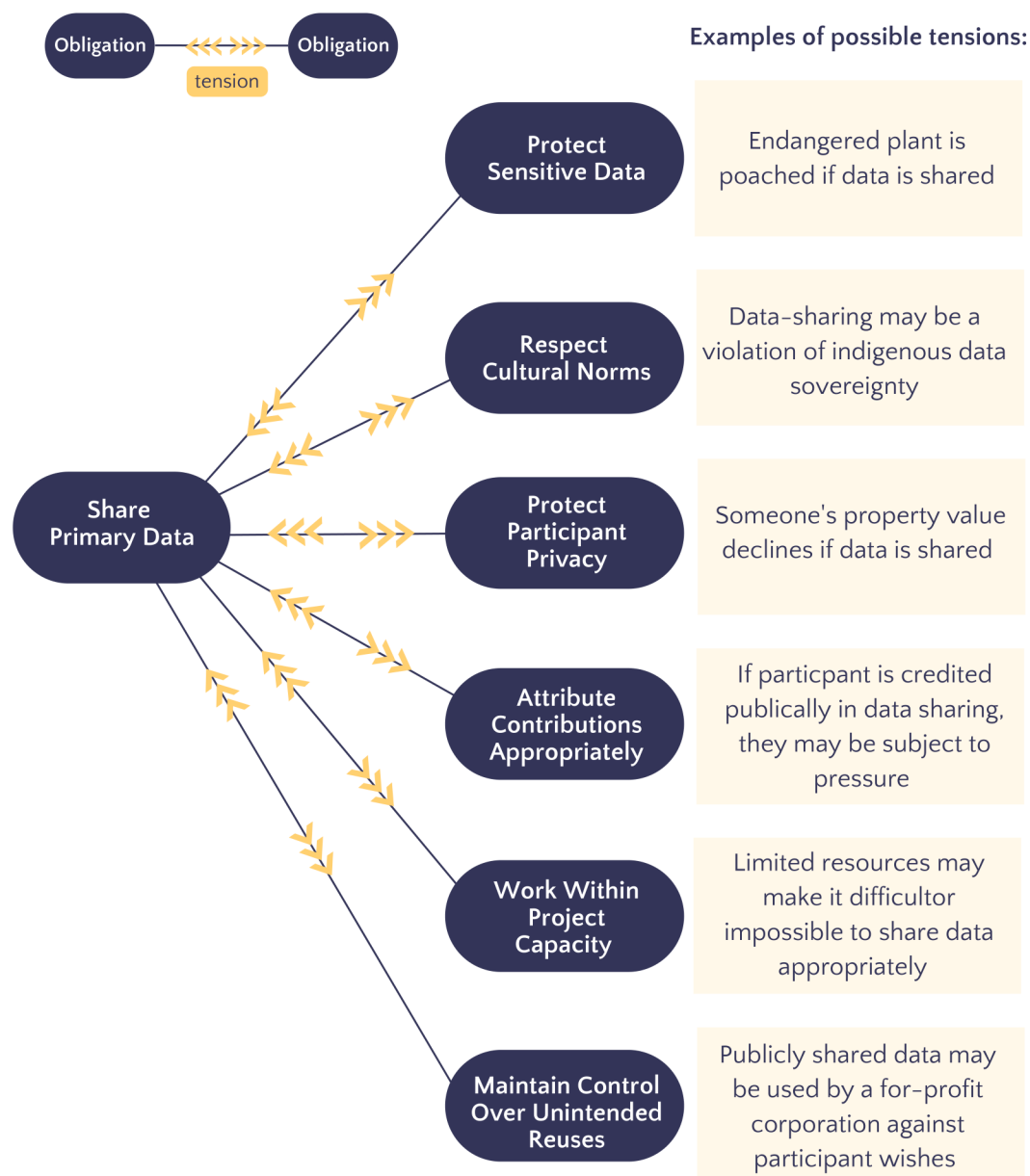
2. RISKS & TENSIONS WITH REPORT-OUTS

Project leaders are data stewards acting on behalf of participants who collectively generated the dataset. To meet the obligation to share data and its documentation with the scientific community, project leaders may need to identify ethical tensions, consider practical realities, and work to ensure that participants understand and endorse the benefits and implications of report-outs. At the same time, project leaders should understand participant perspectives and try to meet their expectations. Thus, they may have responsibilities to assess their data-sharing obligations in the context of other obligations, in particular, the need to protect participant interests.

Balancing conflicting obligations may mean that it is impossible in some circumstances to fully meet all of them. In such cases, it may be worth considering how to mitigate the consequences of that inability, such as through transparency in decisions and communications about them.

The following is an example of potential ethical obligations and tensions in data sharing. You should consider which are relevant to your project so you can develop a plan for addressing them.

Examples of Tensions Between Ethical Obligations in Report-Outs



3. PREPARING REPORT-OUTS

Who delivers report-outs?

Project leaders with access to contributed data are responsible for stewarding data and executing data-sharing decisions. Creating report-outs requires resources like time, knowledge, technical skills, funding, and platforms where data can be securely stored and accessed.

Who develops report-outs?

The decisions you made about governance structure, and reflections on positionality, in the Data Governance Worksheet will help you determine who should be involved in developing a plan for report-outs. These decisions should be informed by an understanding of project objectives, outcomes, and the risks and benefits of making certain data accessible.

Who uses report-outs?

Potential users of report-outs include scientists, researchers, activists, teachers, lawyers, policymakers, members of the public, and a project's own participants and partners. However, technical choices about the form of a report-out (such as data licensing) may limit who is able to make use of a specific report-out.

How will report-outs be accessed?

The main choice when reporting out data is determining how widely and easily they can be accessed and used. Placing data in the private domain restricts access, while placing them in a public domain means the data effectively belongs to everyone and no one in particular.



Pause and Think

What actions do project leaders need to take to make data available in the public domain

Tips:

Data must be explicitly licensed to be part of the public domain. Without a license, the default is that data are private.

Data sharing requires choosing a domain (public or private) or license, which are legal ways to restrict data use and re-use. The Creative Commons includes a globally recognized suite of designations for more readily sharing datasets. The chart below provides some examples of data use and reuse options through Creative Commons licensing.

Data access options exist within a dynamic legal and regulatory terrain. Project leaders should turn to experts for information about current options. (Refer to the Center for Open Science as a starting point for understanding the options and their implications.)

Examples of Creative Commons Domains

Creative Commons Domain:

Example:

CCO

A dataset with this designation can be used freely by anyone without any attribution, restrictions, permissions, or licensing.

CC-BY

A dataset with this designation can be freely used by anyone as long as they provide attribution of the license holder.

CCO

A dataset with this designation can be freely used by anyone for non-commercial uses only.

Bottom Line

Responsibly making data available for the scientific community's use and reuse with attention to ethical, practical, legal, and technical considerations helps contribute to the positive impact of the participatory sciences.

Report-Outs: Worksheet For Your Project

PART 1: IDENTIFY DECISION-MAKERS & RECIPIENTS

Consider implications of your data governance plan for report-outs. Then consider whether any Report-Out decisions require changes to your data governance plan.

1. How might the decision-makers' positionality cause them to overlook appropriate content and structure of report-outs?

Project leaders as STEM professionals may feel obligations to support data reuse, while participants may feel their data contributions should only be used for their own purposes.

2. Are there reasons to change the governance structure for report-outs, either concentrating or diffusing power over decisions about the content and structure of report-outs?

With shared control governance structure, project leaders ask participants, or representatives of types of participants, about their interests at the start of a project and for its duration. With individualized control, structures could allow participants to decide how/when the data points they contributed are included in report-outs.

3. Are there any changes that need to be made to your report-out approaches?

Participants may prefer some forms of licensing over others.

PART 2: DEVELOP AND REFINE YOUR REPORT-OUTS

A. Establish Foundations

1. What are the project's goals for sharing data with the scientific community?

Disciplinary norms and/or journal requirements may stipulate details about data sharing, data domains, and data licensing to allow for verification or reuse.

2. Will the dataset be in public or private domain?

If the data will be placed in the public domain, it may be freely used without permission from an owner with some or no restrictions.

3. Who will be given credit for and take responsibility of the dataset?

This will likely be the project leader(s) and may also include participants and partners.

4. What protections (e.g., as reflected in the CC domains options) will be placed on the data?

If the dataset is licensed with CC-NC then it cannot be used for commercial purposes.

B. CONSIDER POTENTIAL CONSTRAINTS

5. With the help of the figure showing Ethical Tensions, identify which tensions your report-out plans may encounter and consider whether you need to change your plans to address them.

To mitigate privacy risks, you may choose to limit the types of data shared, obfuscate geolocation data, or choose a more restrictive data license.

6. What expectations do your participants and/or partners have regarding sharing data with the scientific community?

Their expectations - for example, regarding data re-use - may be in tension with the disciplinary norms you identified in 2.A.1 above.

7. To what extent can your project accommodate individualized control of the sharing of the primary data that they generated?

Projects may be able to allow participants to opt in or out of retaining control of their data, meaning those data points might not be available for use or re-use by others outside the project.

8. What practical considerations will affect your report-out formats?

Project finances, personnel time and energy, or even the requirements of funders may help determine the nature of your report-outs.

9. What technical considerations will affect your report-out formats?

If you choose a particular platform to collect and store your data, that may affect how you can report them out to a wider community.

10. Are there any legal considerations that may affect your report-outs?

Proprietary ownership of data or databases, licensing requirements, confidentiality of information, or other issues may restrict what can be reported out.

11. How will you resolve the tensions you identified in the questions above? How might you modify your governance structure to identify and incorporate participant perspectives in these decisions?



BRING IT TOGETHER

Writing out your plan is a good way to consolidate and solidify your ideas. Based on the work you were invited to do above, including iterating and revising as necessary, summarize your project's Report-Outs Plan in Box 3 of the Trustworthy Data Practices Summary.

Recompense: Key Concepts

What is Recompense?

In this Toolkit, “recompense” is the term used to indicate how participants and partners obtain appropriate benefits from their contributions to a project. Ensuring recompense means structuring a project so that participants and partners accrue benefits from the range of ways they engage in the project.



Recompense ensures that participants and partners benefit from their efforts

Why Does Recompense Matter?

Participants and partners are essential to the research process, and their efforts deserve recompense. If they do not accrue benefits, projects risk engaging in extractive research practices, meaning that the interests of researchers are being pursued without attention to the interests of participants.

1. THE ROLE OF RECOMPENSE

Recompense is fundamental to projects because participants are:

Essential:

Without the contributions of participants and partners, many discoveries and other project outcomes would not be possible.

Accountable:

Participants and partners share responsibility with researchers for the trustworthiness of discoveries and other project outcomes.

Credible

By affiliating with, or publicly representing, a project, participants and partners may gain credibility and legitimacy for their efforts.

Experts:

Participants and partners often possess expert knowledge of the environment and/or their community, which is critical to producing reliable and robust knowledge.



Pause and Think

In what ways might participants' perceptions of benefits differ from project leaders?

Tip:

Participants may value public expressions of appreciation while project leaders may pursue publications

2. APPROACHES TO RECOMPENSE

Recompense approaches vary based on project design as well as participant and partner motivations, engagement, and preferences. More than one may be appropriate for your project, and benefits will vary based on the nature and extent of involvement in a project.

Public Recognition

Explicitly and publicly acknowledge participants and/or partners for their efforts.

This may occur via project websites, in policy reports, at public events, or through newsletters, certificates, or project-related paraphernalia.

Academic Recognition

Based on merit, name participants and/or partners in peer-reviewed research publications and presentations. This can include authorship, database citation, acknowledgement, or other credit.

Financial Compensation

Compensate participants and/or partners financially depending on the nature of engagement.

This might include direct payment, usage of project-required equipment, or gift cards.

Enrichment Opportunities

Provide opportunities for participants and partners that enhance their experience, knowledge, and/or skills.

Action Oriented Research

Design projects to advance the interests of participants and partners.

Intellectual Property

Acknowledge intellectual property rights that participants or partners may contribute to a project or help to create through their participation.



Pause and Think

How might recompense options differ with the size of the project?

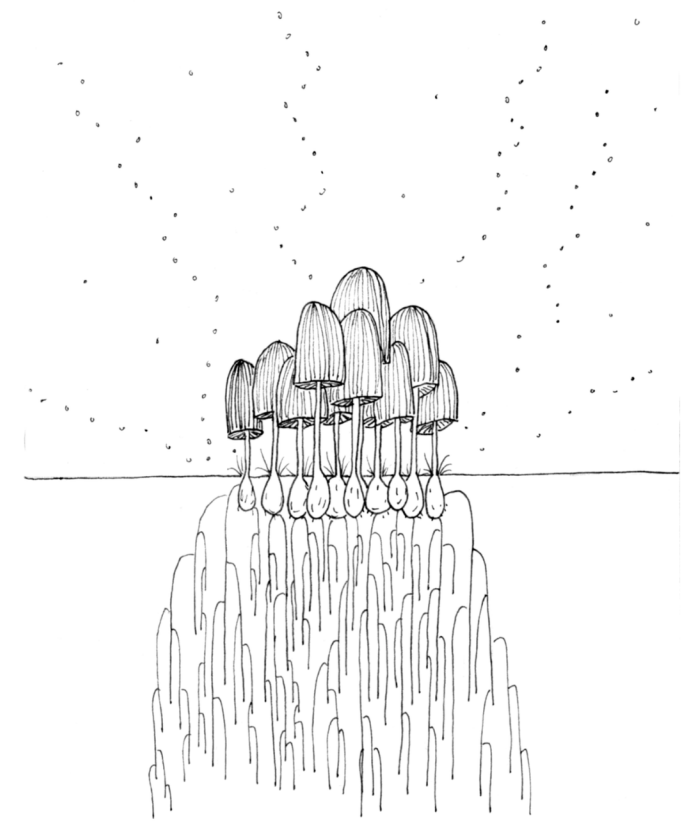
Tips:

A project with 10 participants might have a recompense structure with options for participant authors or payments, while a project with 10,000 participants might have a recompense structure with trainings for professional or personal development.



Examples of Recompense Approaches:

| | Examples: |
|---------------------------------|--|
| Public Recognition | <ul style="list-style-type: none"> • A leader board or special mention in newsletters for highly involved “super user” participants and/or partners • Certificates or badges for participants who take on roles as communicators, educators, or advocates for a project. |
| Academic Recognition | <ul style="list-style-type: none"> • Authorship for those who make a substantial contribution to the research design, data collection, data analysis, and/or preparation of material for dissemination. • Database credit to those who contribute to detailed datasets. • Acknowledgments in publications for those who contribute project data and results. |
| Financial Compensation | <ul style="list-style-type: none"> • Direct payment to those of traditionally marginalized or exploited communities who serve as experts • Human subject compensation may be warranted for those who participate in interventions or contribute personally identifiable information to a research project. |
| Enrichment Opportunities | <ul style="list-style-type: none"> • Personal enjoyment opportunities for those who want to be part of a community working towards a shared goal and/or increase their knowledge about a topic. • Professional development opportunities for those who want to improve their research skills or technical abilities, earn service hours for their resume, or expand their network. |
| Action Oriented Research | <ul style="list-style-type: none"> • Solve problems, such as environmental injustices, rather than only diagnosing them. • Empower communities by equipping participants and/or partners with resources, data, training, certificates, and capacity-building and credibility-building experiences needed to create change. |
| Intellectual Property | <ul style="list-style-type: none"> • Copyright, patents, or trademarks are forms of intellectual property that participants and/or partners may deserve. They may contribute materials they already own • (e.g., photographs) or may help generate intellectual property through project activities (e.g., oral storytelling). |



Bottom Line

Participatory sciences depend on the engagement of individuals and communities. Projects should be structured so that participants and partners share in the equitable distribution of project benefits.

Recompense: Worksheet For Your Project

PART 1: IDENTIFY DECISION-MAKERS AND RECIPIENTS

View your Data Governance Plan as a living document and revise as needed for this tool. Recompense approaches should be prioritized and revisited throughout the life of a project.

1. How might the decision-makers' positionality cause them to overlook appropriate approaches to recompense?

Project leaders may value authorship as the career currency of STEM professionals while participants may value access to resources and credibility to bring about changes.

2. Are there reasons to change the governance structure, either concentrating or diffusing power over decisions about recompense options?

Many recompense options may be necessary if participants are highly heterogeneous.

3. Are there any changes that need to be made to your recompense approaches?

Participants may prefer some forms of recompense over others and should be involved in such decisions.

PART 2: DEVELOP AND REFINE RECOMPENSE OPTIONS

A. Establish Foundations

1. What motivates participants and/or partners to contribute to the project?

Fairness in distribution of benefits involves choosing approach(es) that address both project goals and participant and/or partner motivations.

2. Which of the following approaches to recompense best acknowledge those motivations?

- *Public Recognition*
- *Enrichment Opportunities*
- *Academic recognition*
- *Action-Oriented Research*
- *Financial Compensation*
- *Intellectual Property*
- *Other*

3. What recompense approaches identified above align with your project goals, and are there any approaches that are in conflict or tension?

While rare, it may be possible that limited project resources mean that choosing one type of recompense makes fewer resources available for achieving other project goals.

B. Consider Potential Constraints

1. How might project leader's/team's views on their role as scientist in society constrain recompense options?

Some scientists will not engage in what they perceive to be activism, even if participants request it.

2. Are there professional norms that influence appropriate forms of recompense?

A participant who conducts data analysis may be named an author on a publication, while someone who helps collect data may warrant credit in acknowledgments.

3. To what extent can your project accommodate individual preferences?

It may be appropriate for participants to have a choice in how they are recognized (e.g., anonymously, by pseudonym, by name), to what degree their contributions are known (e.g., data points identified, obscured, or kept private), or to opt-out of recognition entirely.

4. How does the nature and frequency of participation affect recompense options?

Different roles (e.g., data contribution vs. data analysis) or degrees of engagement (i.e., monthly, weekly, one-time contribution, etc.) may warrant different forms of recompense.

5. How does the number of participants affect recompense options?

For example, a small group may be easily acknowledged by name, whereas it could be more feasible to recognize hundreds of participants under a group name instead.

6. Are there other aspects of project design that constrain recompense options?

Anonymous participation to protect privacy will constrain the ability to acknowledge contributors by name.



BRING IT TOGETHER

Writing out your plan is a good way to consolidate and solidify your ideas. Based on the work you were invited to do above, including iterating and revising as necessary, and consulting participants and/or partners, summarize your project's recompense plan in Box 4 of the Trustworthy Data Practices Summary. Keep in mind that more than one form of recompense may be appropriate for your project.

Report-Backs: Key Concepts

What Are Report-Backs?

Report-backs are communications from project leaders to participants and partners for the purpose of sharing project data back in ways that are understandable and usable to them.



Report-backs satisfy a reciprocal obligation to participants and partners

Why Do Report-Outs Matter?

Data contributions from participants and partners generate an obligation for project leaders to share results back with them. Report-backs satisfy this ethical obligation to participants and partners, and are different from report-outs, which satisfy an ethical obligation to share data with the scientific community. Report-backs also convey benefits for project participants.

1. BENEFITS OF REPORT-BACKS

Report-backs to project participants have many advantages:

Engagement

Sharing project data back to participants and partners in a way that is accessible, understandable, and usable can enable them to query, examine, and engage with the data themselves.

Satisfaction

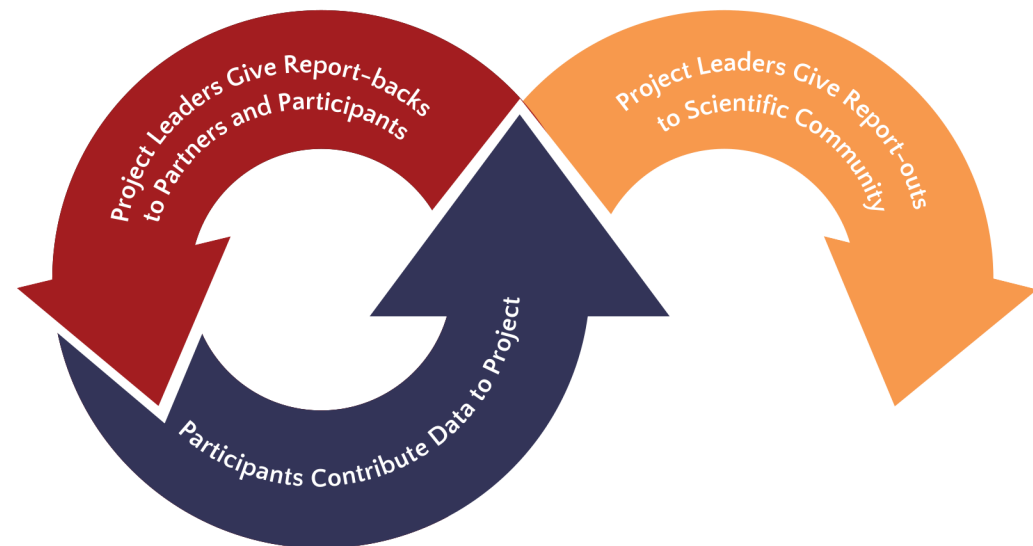
Participants and partners often contribute to research for their enjoyment, sense of belonging, or as a way to make an impact. Report-backs can increase participants and partner satisfaction and retention.

Empowerment

Making meaning of project data for participants and partners can empower people, inspire action, and create new derivative works that go beyond the initial project objectives or capacity.

In addition to the benefits above, report-backs are a way for project leaders to meet obligations to participants, while report-outs meet obligations to the scientific community.

Ways project leaders meet obligations



2. REPORT BACK FORMATS

Data have the most utility if presented so that participants and partners can interpret, understand, learn from, rely on, leverage, and/or use data for their benefit. Participants and partners may appreciate data in a variety of report-back formats.

Raw Data

Original data contribution with documentation

Processed Data

Analyzed, filtered, graphed, or mapped data

Raw and/or processed data may be presented in any of the below summary formats.

Individualized Summary

A report of an individual's own data contribution

Collective Summary

A report of all participant data contributions as a whole

Contextualized Summary

An individual's data within context of other participant data

Examples of different ways to summarize data in report-backs

| | Example 1: Soil Microbes | Example 2: Bird Nesting |
|-----------------------------------|---|--|
| Individualized Report-Back | <ul style="list-style-type: none"> A report of an individual's soil microbes. <i>E.g., Lab results find that participant Alpha's soil has 3 types of bacteria.</i> | <ul style="list-style-type: none"> An individual's observation of birds initiating nests. <i>E.g., Participant Beta observed a bird starting to nest on March 3rd.</i> |
| Collective Report-Back | <ul style="list-style-type: none"> Biodiversity of microbes from all soil samples submitted by participants <i>E.g., Lab results find an average of 30 types of bacteria in soil samples.</i> | <ul style="list-style-type: none"> Timing of bird nesting based on all observations submitted by participants. <i>E.g., Birds began nesting between March 1st and March 30th.</i> |
| Contextualized Report-Back | <ul style="list-style-type: none"> Biodiversity of microbes in soil samples submitted by all participants compared to an individual sample. <i>E.g., Participant Alpha's soil contained low diversity with 3 types of bacteria in their soil compared to an average of 30 types per sample.</i> | <ul style="list-style-type: none"> Timing of bird nesting based on all observations submitted by participants as compared to an individual's observation. <i>E.g., Participant Beta observed a bird nesting on March 3rd, which is early in the season that spanned birds initiating nesting from March 1 to March 30.</i> |

Bottom Line

Providing report-backs that are understandable and usable to participants and partners recognizes them as essential contributors to and beneficiaries of the research process, and satisfies an ethical obligation to them.

> Report Backs: Worksheet For Your Project

PART 1: IDENTIFY DECISION-MAKERS AND RECIPIENTS

To meet the obligation to report back data to participants and partners, project leaders may need to establish what information contributors want by asking about or anticipating their preferences. View your Data Governance Plan and revise as needed for this tool.

1. How might the decision-makers' positionality cause them to overlook appropriate content and style for report-backs?

2. Are there reasons to change the governance structure to concentrate or diffuse decision-making?

Shared governance and individualized control could both be ways to design report-backs that meet the needs and interests of participants.

3. Are there any changes that need to be made to your report-back plan?

PART 2: DEVELOP AND REFINE YOUR REPORT-BACKS

A. Establish Foundations

1. What motivates participants and/or partners to contribute to the project?

If participants join a project for community-building, they may appreciate report-backs related to how their contribution helps achieve project goals. If it is for environmental or social action, they may want access to data and recommendations for change.

2. What do participants and/or partners expect from report-backs in the context of the data they contribute?

Some data may need further processing by project leaders (i.e., DNA samples) or for context (i.e., distribution of species observations) to be meaningful.

3. What can participants and partners learn from report-backs?

Project leaders should be clear about what can be learned from report-backs, especially in cases where it is difficult to provide enough background to avoid misunderstandings and inappropriate conclusions.

4. How will report-backs be best communicated to your participants and/or partners?

Complex report-backs can contain data visualizations, infographics, and data story maps.

5. What report-back formats will convey the benefits desired?

A contextualized report-back might enable a participant to advocate for local policy changes, while a collective report-back may be useful for learning about local animal behavior.

B. Consider Potential Constraints

1. What is the projected time frame for report-backs and how often will report-backs be made?

Report-backs may not be immediate if data collection and analysis require significant time. It may also be necessary to share back data periodically rather than waiting until the project is completed.

2. What resource constraints affect your project's choice of report-backs?

Report-backs may be challenging because of limited resources, energy, funding, skills, or competing priorities.

3. How does the number of participants affect report-back options?

Large participant numbers may mean that some report-back options are not feasible.

4. What potential harms my result from report-backs?

Sharing sensitive information (whether enrollment data or primary project data) may expose participant information, generate tensions among participants, and/or be circulated beyond the participant community and used out of context.

5. Given the above, what measures will you take in developing report-backs to maximize benefits and minimize harms?

Report-backs may need to balance a tension between privacy (e.g., do not disclose sensitive information about harmful infrastructures) and justice (e.g., disclosing location of harmful infrastructure so it can be improved).



BRING IT TOGETHER

Writing out your plan is a good way to consolidate and solidify your ideas. Based on the work you were invited to do above, including iterating and revising as necessary, and consulting participants and/or partners, summarize your project's report-back plan in Box 5 of the Trustworthy Data Practices Summary.. Keep in mind that more than one form of report-back may be appropriate for your project.

Trustworthy Data Practices Summary



BRING IT TOGETHER

This summary is a place to bring it all together. Below you should record the decisions you've made about your project as guided by the individual worksheets, including iterating and revising as necessary.

1: DATA GOVERNANCE

Project leaders are data stewards accountable for executing data governance decisions. To identify and address ethical issues in data governance, project leaders should establish governance structures to diffuse the concentration of power and control. Your positionality and governance structure(s) play a central role in considering ethical obligations in your project.

Plan:

2: DATA INTEGRITY

Trustworthy data are a cornerstone of participatory sciences and dependent on integrity in data generation. Project leaders and participants share the responsibility for data integrity, and leaders are also obligated to ensure that projects are structured to foster and support data integrity.

Plan:

3. REPORT-OUTS

Responsibly making data available for the scientific community's use and reuse with attention to ethical, practical, legal, and technical considerations helps contribute to the positive impact of the participatory sciences.

Plan:

5. REPORT-BACKS

Providing report-backs that are understandable and usable to participants and partners recognizes them as essential contributors to and beneficiaries of the research process, and satisfies an ethical obligation to them.

Plan:

4. RECOMPENSE

Participatory sciences depend on the engagement of individuals and communities. Projects should be structured so that participants and partners share in the equitable distribution of project benefits.

Plan:

The decisions you have compiled above summarize your trustworthy data practice plan. Planning is an iterative practice, so whenever you make changes in your project, you should revisit your plan. You should also think about how to put your plan in lay terms and communicate it to project participants and partners via websites, newsletters, emails, or whatever is appropriate for your project.

CSA wants to maintain these norms of trustworthy data practices, so on CSA Connect, there's a place to share your plan with the community of practice.