

ETHICALLY MAKING & SHARING DATA

A zine for researchers or anyone else!



Learn about the data
life cycle

Review data ethics
principles & precedents



Identify key actions at
every step of the cycle

Put it into practice with
tools & resources



More at <https://bit.ly/data-ethics-rdds!>

Principles & Precedents

“The Belmont Report” on Ethical Principles and Guidelines for the Protection of Human Subjects of Research (1979)

- Published in response to public and congressional concern about unethical research practices (for instance, the Tuskegee Experiment)
- Lays out principles that are now implemented by Institutional Review Boards (IRBs)



Beneficence

The commitment to maximize benefits and avoid causing harm to the extent possible, even if it is not a formal or legal requirement



Respect for Persons

The responsibility to uphold people's power to make decisions that are in their best interest and to protect people who do not have that power



Justice

The commitment to the fair distribution of burdens and benefits among people

Data collection

Centering ethics in the **Acquisition** phase of the data life cycle

Definition

The stage at which people decide which data to collect and why (**conception**), determine how to collect them (**instrumentation**), and take action to obtain them (**collection**).

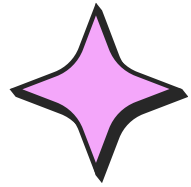
Actions

- Involve stakeholders from the start
- Collect only the minimum viable data
- Require informed consent and “opt-in”
- Provide reciprocal value in exchange for data
- Anticipate potential security risks, such as re-identification

TIP!

Be aware that cloud data storage systems (including Google Drive and Microsoft 360) may harvest your data for AI training without your knowledge!

Data analysis



Centering ethics in the **Processing & Analysis** phase of the data life cycle

Definition

The stage at which people decide how to **tabulate** and **interpret** data. This includes determining which data to include or exclude in analyses, how to process the data to create new variables or measures like indexes, and which people will be involved in the process to decide what the data mean.

Actions

- Forefront the human element inherent in data, rather than approach it as an abstraction
- Be transparent about the data's context and limitations
- Seek out and incorporate stakeholders' interpretation of the data
- Identify encoded bias and consider mitigation actions

Data sharing



Centering ethics in the
Dissemination phase of the data
life cycle

Definition

The stage at which people publish and share data and/or report findings from their analysis.

Actions

- Incorporate quality control frameworks such as FAIR, CARE, and domain-specific specifications
- Account for how publication may reinforce inequities or close disparities
- Share to reduce the burden of duplicate data collection
- Return data and research results to stakeholders and contributors in a form they can use



RDDS staff can help UPenn affiliates assess available options for data sharing based on their needs. Contact us at libraryrdds@pobox.upenn.edu.

Data afterlives



Centering ethics in the **Disposition** phase of the data life cycle

Definition

The stage at which people **destroy** or **archive** data, either completely or partially.

Actions

- Empower individuals to order the destruction of their data
- Be transparent about the plans for the data after project conclusion
- Put in place accountability mechanisms to provide redress for harms that may arise from data misuse

Tools & Resources

Learn more

- D'Ignazio, Catherine, and Lauren F. Klein. *Data Feminism*. Cambridge, MA: The MIT Press, 2020.
- D'Ignazio, Catherine, and Lauren F. Klein. "Data Feminism for AI," *FACcT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*. Rio de Janeiro, Brazil: June 5, 2024.
<https://doi.org/10.1145/3630106.3658543>
- Benjamin, Ruha. *Race After Technology: Abolitionist Tools for the New Jim Code*. Oxford, England: Polity, 2019.
- data.org resources, courses, and playbooks

Put it into practice

- [ODI Interactive Data Ethics Canvas](#)
- [Data Ethics Toolkit for the Participatory Sciences](#) (AAPS)
- Grantees may offer additional guidance, like the [Packard Foundation Data Ethics Guidebook and Toolkit](#)

ABOUT THIS ZINE

This zine created by Cynthia Heider, based on a workshop offered at Penn Libraries.

- Download it at <https://bit.ly/data-ethics-rdds>
- Feel free to print and distribute under the terms of [CC BY-NC-SA](#).
- Preferred citation: Heider, Cynthia. "Ethically Making and Sharing Data," accessed [insert date you accessed it here], https://github.com/upenndigitalscholarship/workshop-data-ethics/blob/main/2025-Zine_Ethically_Making_Sharing_Data.pdf.



To click through the links in this zine, view it digitally at [https://bit.ly/data-ethics-rdds!](https://bit.ly/data-ethics-rdds)



Penn Libraries

UNIVERSITY *of* PENNSYLVANIA

Last updated February 12, 2025.