Ethical & Responsible Data Management

Renata Curty
Research Data Services, UCSB Library
rds@library.ucsb.edu

UC **SANTA BARBARA**

Learning Objectives

- Understand core ethical and responsible data management principles, particularly in the areas of self-governance, managing conflicts of interest, mitigating bias and harm, and appropriately handling sensitive data.

- Articulate these principles in the context of real-world scenarios relevant to environmental sciences research.

Data Management

- Requires an integrative view of the project and the whole data lifecycle

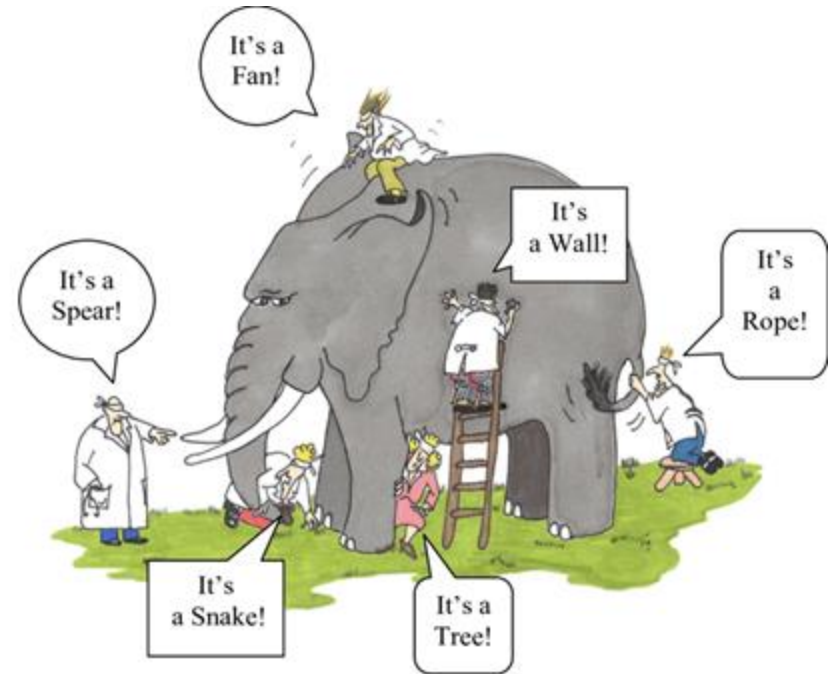- Encompasses **legal, socio-technical and ethical** aspects

Image: Renee Guzlas

Data Management Roles

Data Collector • Instrumentation Specialist • Metadata Creator

Data Analyst • Database Manager • Data Curator

Ethics? Whose responsibility?

Image: Renee Guzlas

Ethics

Standards or agreements of right or wrong, that morally (and often legally), orients behavior.

Environmental Sciences

Special **interconnectedness of human, ecological, and planetary well-being**, as well as the **long-term and global consequences** of actions in the field.

Group Activity - Reflection

- What ethical concerns did you encounter during your capstone project (data, data provider, client, or community of interest)?

- How did you address or mitigate these ethical concerns throughout the project?

bit.ly/eds213-week8

The Ecology of Data Ethics

- Mirrors a ecosystem, where data flows through and is shaped by people, organizations, technologies, and networks.

- Rather than viewing data management solely as a technical process, emphasizes how human values and decisions shape data's impact and meaning.

- Researchers are "residents" within the data ecosystem, both influencing and being influenced by the broader data environment.

Autonomy | Fairness | Well-being

See: https://sites.psu.edu/ethicsofdatamanagement

Key Guiding Principles

Autonomy **self-determination and governance**

Fairness **equitable and just treatment**

*Well-being* **promote positive outcomes and minimize harm**

Autonomy

Dawson, N. M., Coolsaet, B., Bhardwaj, A., Booker, F., Brown, D., Lliso, B., Loos, J., Martin, A., Oliva, M., Pascual, U., Sherpa, P., & Worsdell, T. (2024). Is it just conservation? A typology of Indigenous peoples' and local communities' roles in conserving biodiversity. *One Earth, 7*(6), 1007–1021. https://doi.org/10.1016/j.oneear.2024.05.001

Governing Indigenous Data Across the Data Lifecycle
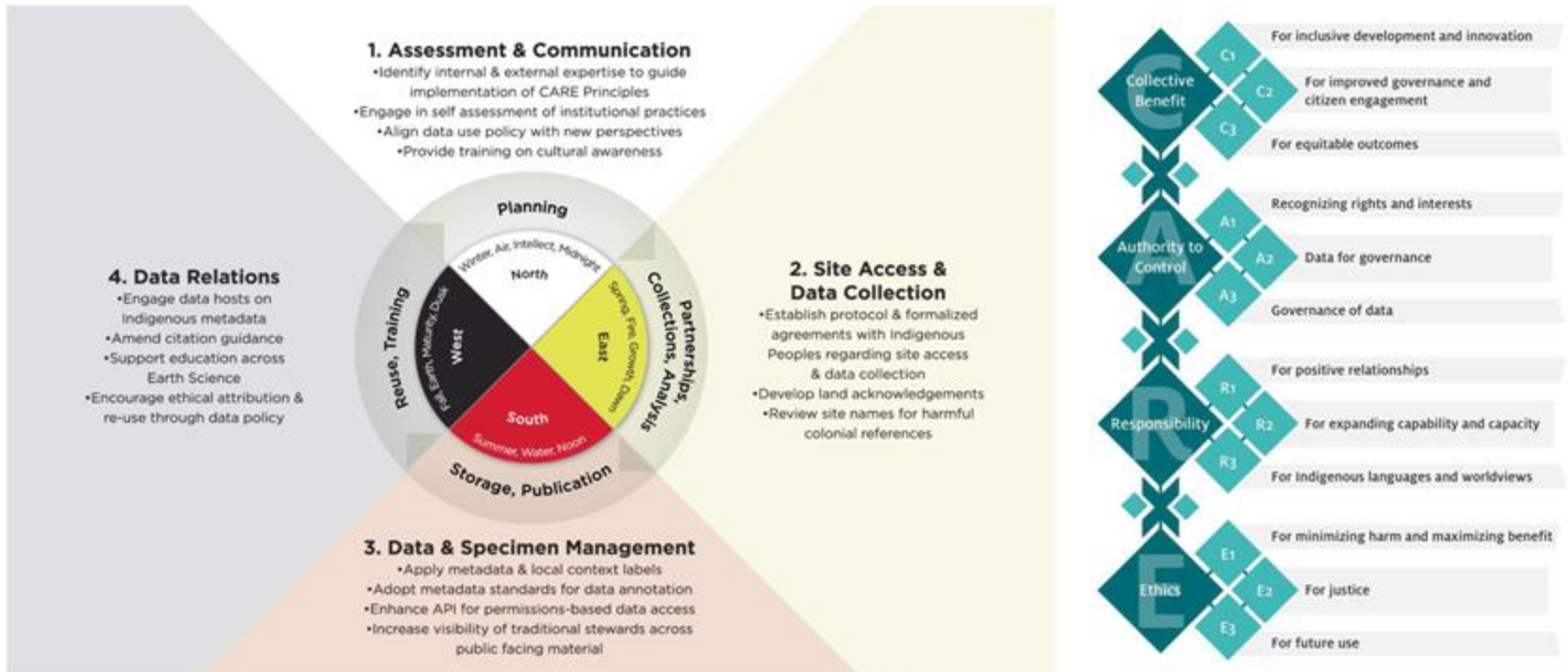
Jennings, L., Jones, K., Taitingfong, R. et al. Governance of Indigenous data in open earth systems science. *Nat Commun* 16, 572 (2025). https://doi.org/10.1038/s41467-024-53480-2

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., ... & Hudson, M. (2023). The CARE principles for indigenous data governance. *Open Scholarship Press Curated Volumes: Policy*. https://doi.org/10.5334/dsj-2020-043

Impact on DM

- Joint and informed decisions
- Consider Ongoing Consent (when possible)
  - Waves of data collection
  - Shifts on scope or priorities
  - Renegotiate terms
- Data ownership and access control!
  - Who can access the data?
  - How it can be used or shared?

"'Accessible' means somebody can access it but you are still allowed to regulate it."
Rosie Alegado
University of Hawaii

Ravindran, S. (2024). Open with care. *Science*, *386*(6720), 372-375.
https://doi.org/10.1126/science.adu0429

Positionality

Researchers, policymakers, or consultants <u>must declare any competing interests</u> or obligations that could compromise their objectivity or integrity when engaging in research activities.

FOOD FOR THOUGHT

Fisheries Scientist Under Fire For Undisclosed Seafood Industry Funding

MAY 12, 2016 · 7:02 PM ET

By Clare Leschin-Hoar

A fishing dragger hauls in a net full of Atlantic cod, yellowtail flounder and American lobster off the coast of New England. Greenpeace says Ray Hilborn, a prominent fisheries scientist known for challenging studies that show declines in fish populations, failed to fully disclose industry funding on some of his scientific papers.

Jeff Rotman/Getty Images

Greenpeace, for example, alleges the researcher did not disclose $58,000 in funding from the New Zealand Seafood Industry Council in a 2006 paper for the Canadian Journal of Fisheries and Aquatic Sciences on orange roughy. They also say he neglected to mention employment from the California Fisheries Coalition, which includes 27 recreational and commercial fishing associations, while questioning the environmental benefit of establishing marine protected areas off the coast of California.

Source: https://www.npr.org/sections/thesalt/2016/05/12/477827180/fisheries-scientist-becomes-latest-target-of-activist-s-records-requests?utm_source=chatgpt.com

12

Conflict of Interest

Schäffer, A., Groh, K. J., Sigmund, G., Azoulay, D., Backhaus, T., Bertram, M. G., ... & Scheringer, M. (2023). Conflicts of interest in the assessment of chemicals, waste, and pollution. *Environmental science & technology*, 57(48), 19066-19077. https://doi.org/10.1021/acs.est.3c04213

- Improve oversight of peer review contracts by requiring disclosure of industry ties and client relationships from potential panelists.

- Provide additional guidance and training for peer review leaders on evaluating panelists' viewpoints to ensure independence and balance.

13

Fairness
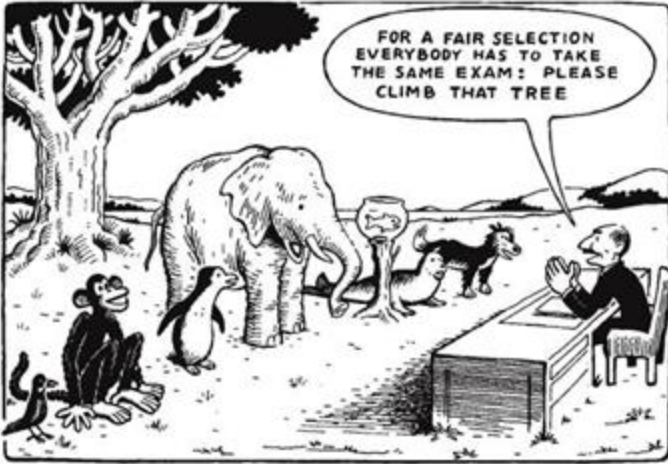
Injustice and Bias Prevention

Prevent negative direct or indirect impact on individuals, groups or communities due to the lack of diversity and representation, existence of bias (unjustifiable concentration on a particular side), or unfairness (discriminating treatment of data and people) present in the data.

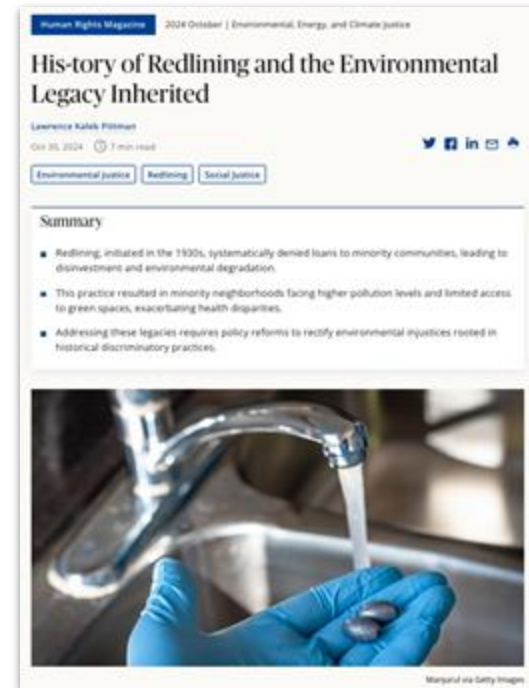Researchers should be aware of power imbalances in the research process.

https://www.sciencedirect.com/science/article/abs/pii/S146290112 1001489

Environmental Justice

Mapping inequality: redlining in new deal America:
https://dsl.richmond.edu/panorama/redlining

Word Cloud from the StoryMap, "The lines. that shape our cities."

Pittman, L. (2024). His-tory of Redlining and the
Environmental Legacy Inherited. *Human Rights, 50*(1/2).
https://www.americanbar.org/groups/crsj/resources/human-
rights/2024-october/history-redlining-environmental-legacy

Impact on DM

- Transparency

  - Are limitations and potential biases (data and/or model) documented and openly disclosed?

- Data enrichment (mitigate imbalances)
  - Consider integrating socioeconomic, demographic, and health data to better understand environmental impacts across different groups.


Lower density areas?

(a) Hail reports follow population

(b) Tornado reports follow population
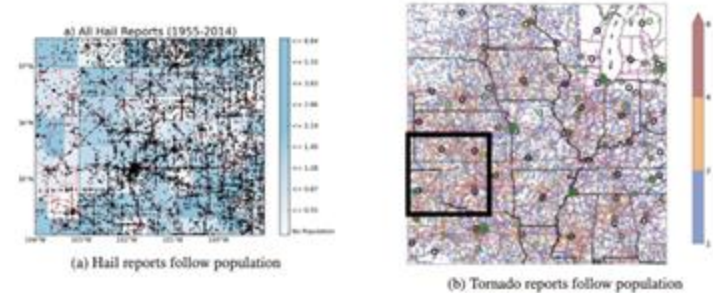
McGovern, A., Ebert-Uphoff, I., Gagne II, D. J., & Bostrom, A. (2022). Why we need to focus on developing ethical, responsible, and trustworthy artificial intelligence approaches for environmental science. *Environmental Data Science*, *1*, e6. https://doi.org/10.1017/eds.2022.5

Well-being

Protect and respect the safety, dignity, choices, and rights of the people, animals, communities, and land involved in or affected by the research.
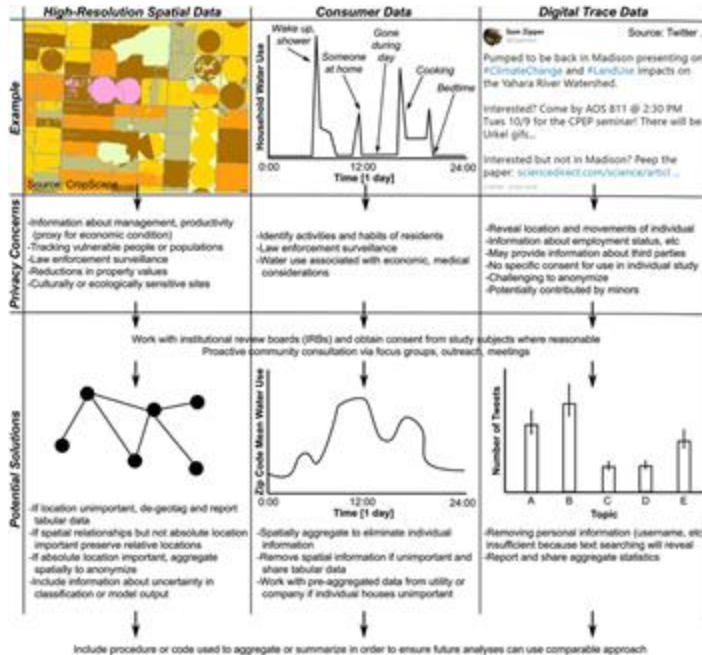
Sensitive Data

Concerns about third parties who could use the data for profit, coercion, or regulatory action

Zipper, S. C., Stack Whitney, K., Deines, J. M., Befus, K. M., Bhatia, U., Albers, S. J., ... & Schlager, E. (2019). Balancing open science and data privacy in the water sciences. *Water Resources Research*, *55*(7), 5202-5211. https://doi.org/10.1029/2019WR025080

Sensitive Data - Beyond Human Subjects!

Four levels of sensitivity (extreme to low) according to **biological significance** and exploitation **threat**.

Key recommendations:

- Generalize the spatial locality or geographic coordinates.

- Do not apply randomization!

Endangered Species

Leatherback Turtle

Threatened or Commercially Exploited Plants

Eelgrass

US. Fishery and Wildlife Service
Endangered Species Library

Protected Research Sites

USGS - PAD-US
Terrestrial and Marine
Protected Areas

Chapman AD (2020). Current Best Practices for Generalizing Sensitive Species Occurrence Data. Copenhagen: GBIF. https://doi.org/10.15468/doc-5jp4-5g10

Impact on DM

- **Data Pre-Processing:** extra cleaning, deid, or transformation steps.
- **Access Control:** who can access data during and after the project.
- **Data Sharing:** Will you share derived/aggregated data or metadata only?
- **Disclosure Risk vs. Utility:** How will you balance privacy protection with data usefulness?
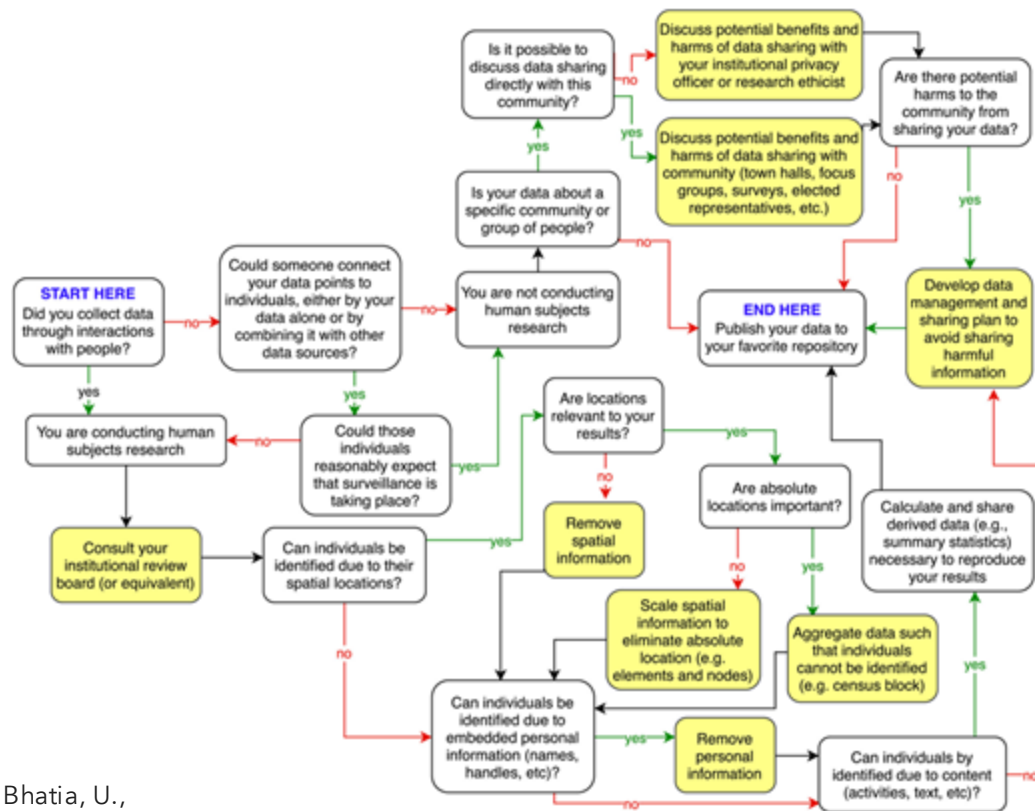
Zipper, S. C., Stack Whitney, K., Deines, J. M., Befus, K. M., Bhatia, U., Albers, S. J., ... & Schlager, E. (2019). Balancing open science and data privacy in the water sciences. *Water Resources Research*, *55*(7), 5202-5211. https://doi.org/10.1029/2019WR025080

19

Data De-identification

The process which removes direct and indirect (quasi) identifiers from data to mitigate privacy risks, while allowing data to be safely shared and reused.

How?

Week 9!
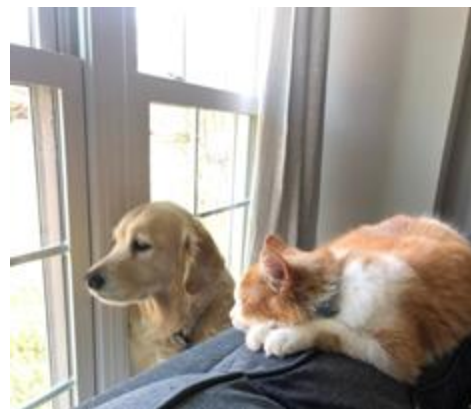
Group Activity - Any Changes?

After reflecting on our discussion and the examples we've reviewed, are there any ethical considerations or lessons you hadn't initially thought of but now realize are relevant to your project?

bit.ly/eds213-week8