EnDEAR:

Environmental Data Ethics and Accessibility Resources

Nastassia Barber, Sarah Buckingham, Alex Powell, Emelia Williams

Why does it matter ?

As data scientists, it is important to recognize the risks and vulnerabilities our research exposes individuals to and mitigate potential harm to the best of our ability by committing to responsible data accessibility and transparency practices.

There are numerous use cases where sharing environmental data can create new risks for communities connected to those environments.

On the other hand, sharing data is extremely useful for advancing both future science and public understanding, and can inform vulnerable groups of unknown risks.

São Paulo water crisis

Fishing communities

Flooding data

What is your solution?

EnDEAR

(Environmental Data Ethics and Accessibility Resources)

An interactive framework for researchers with an inventory of questions and resources to examine the ethics of sharing and communicating your environmental research data

It considers the **critical questions** that need to be asked at each stage of research and provides **resources** to address the many potentialities

Pathway - How do you plan to do this ?

- We've developed a framework of questions related to sharing and accessibility and compiled a number of resources
- We created a GitHub where's hosted
- Need: to add more resources, update based on user interaction

Questions like:

Who could benefit from using this data?

Where will you host the data/in what format?

Does publishing the data create any liability for stakeholders? What are the power dynamics between stakeholders?

What social governance systems are already in place in this community?

Is there any personal information in this data?

Output - What is the expected product ?

- GitHub doc for now
  - Allows people to contribute
  - [Link](#)

- Could move to a different location with more interactivity in the future

---

Environmental Data Ethics and Accessibility Resources (EnDEAR)

What is this about?

This resource was produced as part of the 2024 [NCEAS Environmental Data Science Summit](#). It is intended to help researchers/anyone who produces environmental data generate a series of questions to inform their decision-making about data sharing and communication. Data sharing presents a huge variety of accessibility, privacy, technical, and ethics challenges, and this is far from an exhaustive list. The questions are arranged into three sections depending on when they would ideally be addressed-- Pre-Research, During Research, and Pre-Publication. At the bottom there is also a list of more general resources.

How can I help?

This document is designed to be ever-evolving. Feel free to create a PR to add to this document if you know of any additional resources. This document relies on people with a variety of backgrounds contributing their knowledge. If you have any questions about how to do this, feel free to email nastassiabot@gmail.com.

Stage 1: Pre-research

Why share?

- Who could also make use of this data?
  - Could it be reused in other research projects?
  - Could it be useful or informative to the general public?
- Do stakeholders want this data to be shared?
  - Why or why not?

Accountability

- Who owns this data "legally"? / "ethically"?
  - How do you plan to determine this?
- What intention are you going into research with?
- What community/stakeholders are affected by this research/data sharing?
- What social governance systems are already in place in this community?

Resources:

[Data Governance Institute Data Ownership](#)

Assessing Risk

- Is there any personal information in this data?
  - If so, what level of anonymization or de-identification is needed

Outcome - What is the expected impact ?

Environmental data can invite both risk and opportunity

Developing reusable and adaptable workflows related to data ethics supports effective research and supports communities

Foster conversations to make the best possible decisions around sharing data
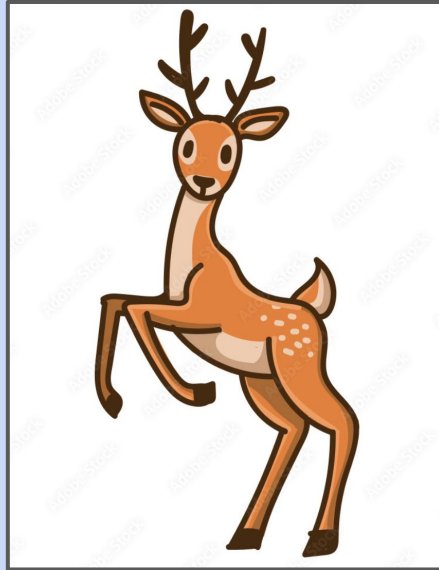
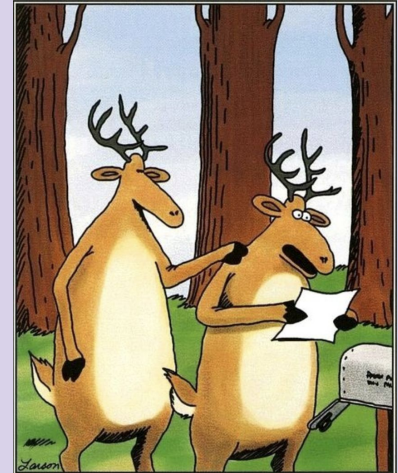Globeville ASARCO Superfund Site research use case

What is needed next ?

Continue compiling resources and research

Continue adding use cases

Edit questions based on user interactions and reactions

Thank you!

Questions?