Nina Di Cara
MRC Integrative Epidemiology Unit

ORCID: 0000-0002-6179-1067

jointly with

Natalie Zelenka

Jean Golding Institute

ORCID: 0000-0002-1007-0286

Data Hazards: collaborative ethical thinking for data science projects

submitted for the category of

Widening Reach / Improving Quality

in the

Open Research Prize 2022

SUMMARY (100 word abstract)

The Data Hazards project aims to create an accessible vocabulary of ethical concerns in data science, that can be used by researchers to prompt thinking about ethics, to discuss ethics across interdisciplinary boundaries and to share worst case scenarios of data science projects. All of these aims intend to increase the quality of ethical thinking in data science, and increase the openness of conversations about potential impacts, so that we can work to prevent data science from doing harm. The Data Hazards and their associated materials have been developed and shared openly with the research community.

What did you do?

Data Hazards is a project to find a shared vocabulary for talking about worst-case scenarios of data science and to use that vocabulary to help people understand and avoid Data Hazards. We created the project as well as materials and workshop templates for other people to use the Data Hazards to improve the consideration of the ethical implications of data science, and to communicate these with other researchers and the public. We hope that this will contribute to an improvement in the quality of data science work, and contribute to preventing future harm.

At their core the Data Hazards are a series of (currently) 11 labels (Figure 1) that each describe a potential risk of a data science project. We made them as images so that they would be visually striking and could be shared by researchers as part of papers, presentation slides or other research outputs. As well as the labels themselves we have developed a reflective workshop format where researchers can get feedback on the impacts of their projects from a diverse group of people in order to expand their thinking and expose them to risks that they may not have thought of. The reflective workshop structure can be adapted to be a lab group meeting, a workshop, a public engagement exercise or a teaching activity, and we provide sample materials to support varied uses.

Our whole project has been developed on a public repository and all of our materials have been published under an open license to allow people to re-use and edit them for their purposes.

Why did you do it?

Data science projects typically do not meet the threshold for ethical review, but they can have huge societal impacts that reach many people's lives. We wanted to create a series of prompts that made it very easy to consider how data ethics might apply to any project, including at the early stages of research when impacts may seem far away. We also wanted to help researchers to work with people from other fields or the public to think of new perspectives on their projects that they might not come across otherwise. As well as hoping to improve the quality of ethical consideration in data science, we wanted people to make their reasoning open and to communicate it with others, with the aim of sharing responsibility for preventing worst-case outcomes as a research community.

How did you do it?

To develop the materials, we ran a test workshop with 25 participants where three researchers presented projects and the other attendees discussed them. They all then gave us feedback on the labels and the exercise format, with our analysis plan pre-registered on the Open Science Framework. Following on from the workshop we then developed a format for using the Data Hazards as teaching materials and worked with lecturers to test these with students.

All our materials are published and developed openly on our public repository/website. Our website also features a public timeline of the project that includes milestones we are working towards and planned future releases of materials. We similarly share all slides from our talks and other outreach work on this website.

What barriers / challenges did you have to overcome?

Both of us have developed this project alongside our usual research or work responsibilities, and so one of our main barriers has been giving as much time to it as we wanted to be able to. However, it is something that we both feel passionate about and so we have made as much time available as we can.

Another challenge has been how to appropriately credit people in an open source project, so that we can recognise the contribution of people who have provided feedback, helped at workshops or told us about how they have made use of the materials. We now use a tool to record all our contributors, and recently introduced a detailed 'Contributors' page where anyone who has been part of the project can write about their contributions in more detail.

What does it mean for you and your research?

The project itself, and the open way in which we have run it, has enabled us to build networks with researchers in other fields and countries who share our aim to make data ethics a bigger priority in data science. This has led to new routes for our work, and new possibilities for making a positive impact, such as now working to train others to run Data Hazards workshops.

How might your findings / approach help other researchers?

It has been exciting for us to see people already starting to use the Data Hazards in their research and teaching based on the open materials that we have made freely available on our website. Our resources have been used by colleagues in Bristol across multiple subjects, and by a lecturer in the United States who found our materials online.

In general, we hope that the Data Hazards themselves encourage a more transparent and thoughtful approach to ethics in data science by encouraging discussion of potential hazardous outcomes of data science across interdisciplinary boundaries. We also hope that using the reflective approach will encourage researchers to see ethical issues from other people's points of view, and to illicit feedback from those outside their immediate environment when thinking about the possible impacts of their work.

Additional Information

Project website: https://very-good-science.github.io/data-hazards/

OSF Pre-registration: https://osf.io/pcv7j

Figure 1: The Data Hazard labels