

Data Hazards, Ethics and Reproducibility One-Day Symposium

In person location: Enigma 2.0 Alan Turing Institute

Online location: Join Zoom Meeting

<https://us02web.zoom.us/j/89670289137?pwd=U2FkQjY0cW9ZakZDQm5RYjV1bnZqZz09>

Meeting ID: 896 7028 9137

Passcode: 313561

Time (GMT)	Topic	Materials	Zoom Link
10:00 - 10:15 AM	Welcome and Introduction to the day	slides	See top of document
10:15 - 11:05	<p>Keynote Speakers (Chair, Susana):</p> <p>Anne Lee Steele - <i>Talk title: From culture to computation: mapping my open research journey</i></p> <p>Anne is the Community Manager for The Turing Way project at The Alan Turing Institute, where she facilitates a collaborative resource for reproducible data science. She is a current fellow at the Internet Society, where she investigates low earth orbit (LEO) satellite connectivity, and was previously a resident at Wikimedia Deutschland's Unlock Accelerator. She holds degrees from Columbia University and The Graduate Institute of International and Development Studies in Geneva, both in anthropology and sociology.</p> <p>Paz Bernaldo - <i>Talk title: Am I in or am I out? Investigating who is in, in open science</i></p> <p>Many in the south(s) of the world grow up in our early years enchanted by nature and the universe and all things feeding scientific mindsets and exploration. But then comes school, and social inequalities, gender disparities, stereotypes and socio</p>	Slides tbc.	

	<p>economic injustices that separate those who can continue to explore and those who can't. My long term project Vuela has been an attempt to question who gets to participate in open science and in what capacity. And in a similar line, OLS represents an opportunity to understand common challenges and windows of opportunity to move away from the ivory tower science that fails to face the challenges of our century.</p> <p>Paz is a key coordinator at the Open Life Sciences project. She has worked on a variety of projects, including a project called Vuela, which questioned who gets to participate in scientific processes and in what capacity, who gets to be called "contributor", and who gets to decide what the problems are in the first place.</p>		
11:05 - 11:15	Break		
11.15 - 13:00	<p>Data Hazards Workshop (Hosting: Susana and Vanessa, helping facilitate online: Nina, Natalie, Eirini)</p> <p>Data Hazards is a project about worst-case scenarios of Data Science. Data Scientists are great at selling our work, for example communicating the gains in efficiency and accuracy, but we are less well-practiced in thinking about the ethical implications of our work. The ethical implications go beyond most ethics Institutional Review Boards, to questions about the wider societal impact of Data Science and algorithms work.</p>	Slides.	
13:00 - 14:00	Lunch break		
14:00 - 14:30	<p>Networking discussion: why do you care about data ethics? (online only). (Online Facilitating: Nina, Patricia, Vanessa)</p>	Slides	

<p>14:30 - 15:00</p>	<p>Reproducibility in PhDs (Chair, Maryam)</p> <p>14.30 - 14.45 Ezra Herman - <i>Talk Title: A reproducible thesis - writing code and reports in one go with Snakemake and R Markdown.</i></p> <p>As data analysis projects come to an end, a report needs to be written. This may be for a supervisory meeting, a thesis chapter or a journal submission. In all these cases, analysts can go through a long back-and-forth, in which code is tweaked and outputs within the report are updated. What if the analysis and the write-up could be merged into one workflow, saving us time and energy? In this talk I will introduce how I use R Markdown and the Snakemake workflow manager to write reproducible code and reports. I will also touch on R Markdown extensions which I have not yet used, but which could make your manuscript or thesis writing easier.</p> <p>14.45 - 15.00 Natalie Zelenka - <i>Talk title: How I tricked myself into writing my thesis (by making it as ethical and reproducible as I could).</i></p> <p>A whistle stop tour of how and why I tricked myself into writing my thesis. I'll talk about what worked well and what didn't as I tried to make a reproducible thesis for a project that contained multiple programming languages (R, Python, bash) using JupyterBook, including interactive images (Plotly), my software documentation (sphinx), automatically running tests and build the thesis website (GitHub Actions), including asides about the worst people in the history of statistics, and illustrating my favourite parts of molecular biology - and a new version of the University of Bristol crest.</p>	<p>Slides tbc.</p>	
<p>15.00 - 15.15</p>	<p>Break</p>		

<p>15:15 - 16:00</p>	<p>Embedding Ethics and Reproducibility into your Research Career (Chair, Roxana).</p> <p>15.15 - 15.30: Alden Conner - <i>Talk Title: <u>The Turing Way: A collaborative guide to data science and research</u></i></p> <p>This talk will introduce you to The Turing Way, which is an open, collaborative resource created as a handbook for reproducible research. I will describe the history of the Turing Way, how the Turing Way community is working to drive culture change, and how to get involved.</p> <p>15.30 - 15.45 Melanie Stefan - <i>Talk title: <u>The ethical lecture: looking at university teaching through a Data Hazards frame</u></i></p> <p>A frequent part of doing science is teaching science, and often, this takes the form of a lecture. Whilst considerations around ethics are built into the research process, for instance through ethics reviews and committees, the same care does not go into the design and delivery of teaching. Here, I use the "Data Hazards" framework to look at frequent practices around university lectures (including my own) and how we may go about improving them.</p> <p>15.45 - 16.00 Clau Fischer - <i>Talk title: <u>The Turing Commons, Training in AI ethics and responsible research</u></i></p> <p>In this presentation, I will introduce the work we have been doing for the Turing Commons and how we hope to scale our work, but I will also talk about how I got to be doing this job (including a career change from economics to data and AI ethics).</p>	<p>Slides tbc.</p>	
<p>16:00-16:15</p>	<p>Facilitated Discussion: Embedding ethics into your research projects - key takeaways</p>		
<p>16:15 - 16:30</p>	<p>Closing remarks.</p>		

Please fill in feedback:

<https://forms.office.com/r/KmTidDcuiC>

