



University  
of Exeter

# Open Research Awards 2024

## **HASS disciplines prize**

**Aditi Dutta**

### **Title of case study**

#### **Is open and reproducible research possible for online sexism?**

##### Introduction to the case study

My thesis topic is understanding the dynamics of online sexism, focusing on Twitter(X). I am looking at various aspects of approaching the issue of online sexism, experienced by women in politics in the UK. Considering that sexism is very objective, they can be implemented differently by computational models that are specifically designed for understanding nuances in languages, such as the natural language models, all of which has been an integral part of my work. My research aims to bring impact in digital society research, and in bringing gender equality in digital spaces.

##### Detail the practices you have used to support openness/reproducibility

I have regularly contributed to data/code sharing in all of my projects. Upon publication or completion of any of my projects, I make sure to upload the code in online repositories/platforms like Github (<https://github.com/booktrackerGirl>), Kaggle (<https://www.kaggle.com/aditidutta>) and Zenodo (information found in <https://orcid.org/0000-0002-1531-5510>). Through association with research institutions like the Turing Institute who are dedicated to open research, I have participated in their collaborative works to support the same (e.g., <https://www.turing.ac.uk/news/publications/data-study-group-final-report-sustrans>). I have also contributed in promoting submissions in open access journals (co-edited 'The Open Review' journal: <https://www.swdtp.ac.uk/wp-content/uploads/2023/10/The-Open-Review-2023-A.pdf>), and myself submitted in similar open access journals for my work.

I have also contributed to The Turing Way's guide for reproducible and ethical research in social data (<https://book.the-turing-way.org>). As a computational social scientist, I am aware of the challenges of openness and reproducibility of our work, especially for social datasets from online platforms. Therefore, I ensure that I could contribute in helping other researchers who face similar situations like me, by providing guidance through contributions such as this.

##### Outline any barriers or challenges you encountered, and how they were handled

The most common challenge faced is making the data available for reproducibility, without giving away any unique primary or secondary identifiers. Considering that the data is from Twitter(X) which has now revoked its free access to multiple researchers like me, I cannot openly share the data. Hence, to counter that, I have devised ways to anonymise the dataset, such that any mentions of entities would be scrapped off from the uploaded dataset, along with its identifiable metadata. The metadata would be available upon request, but there will be a unique key linking them to the entries in the uploaded dataset.

### What benefits have you identified from your case?

The benefits I have identified in my case are (open) access to resources which were integral for my research, and networking with the right institutions who can help us in learning more about open and reproducible research.

### How do you plan to continue this case moving forward?

I plan to continue contributing to open and reproducible research, at any capacities. Through my mentors and associations with universities and institutions dedicated at open research, I am constantly acquiring knowledge on better practices. In social science, it is difficult to access resources to guide us in our research. I had previously started a blog site (<https://booktrackergirl.github.io/The-Curious-Researcher/>) to note down my experiences and resources which could be useful to others. My goal is to start populating the website after completing my PhD with all the knowledge I had acquired in the process, to help other researchers like myself.

### Conclusions

I hope to share about my acquired knowledge to our research community in order to promote FAIR principles in research practices, to promote more researchers participating in open and reproducible research.

### Links to any materials/resources relevant to the case

<https://github.com/booktrackerGirl>

<https://www.kaggle.com/aditidutta>

<https://orcid.org/0000-0002-1531-5510>

<https://www.turing.ac.uk/news/publications/data-study-group-final-report-sustrans>

<https://openlifesci.org//openseeds/ols-7/>

<https://booktrackergirl.github.io/The-Curious-Researcher/>

<https://book.the-turing-way.org>,

<https://www.swdtp.ac.uk/wp-content/uploads/2023/10/The-Open-Review-2023-A.pdf>