

INTRODUCTION TO ETHICS BEST PRACTICES FOR COMPUTATIONAL SOCIAL SCIENTISTS

Developed by Seliem El-Sayed, MA and Univ.-Prof. Dr. Barbara Prainsack, University of Vienna within the project Digitize! Computational Social Science in Digital and Social Transformation, funded by the Austrian Federal Ministry of Education, Science and Research.

The following infographic provides an introductory overview of best practices in research ethics in computational social science (CSS). The infographic has been designed specifically for researchers who are new to the field, at the early stages of their academic careers, or lacking adequate background knowledge in social science research ethics. It serves primarily as a resource to support scholars in developing ethical research practices and to facilitate the work of ethics committees. The infographic covers a range of topics that are crucial for upholding ethical standards in CSS research, including the granularity of data and avoiding the misinterpretation of research results. It provides researchers with both information and practical guidance to facilitate the integration of ethical principles into their research practices.

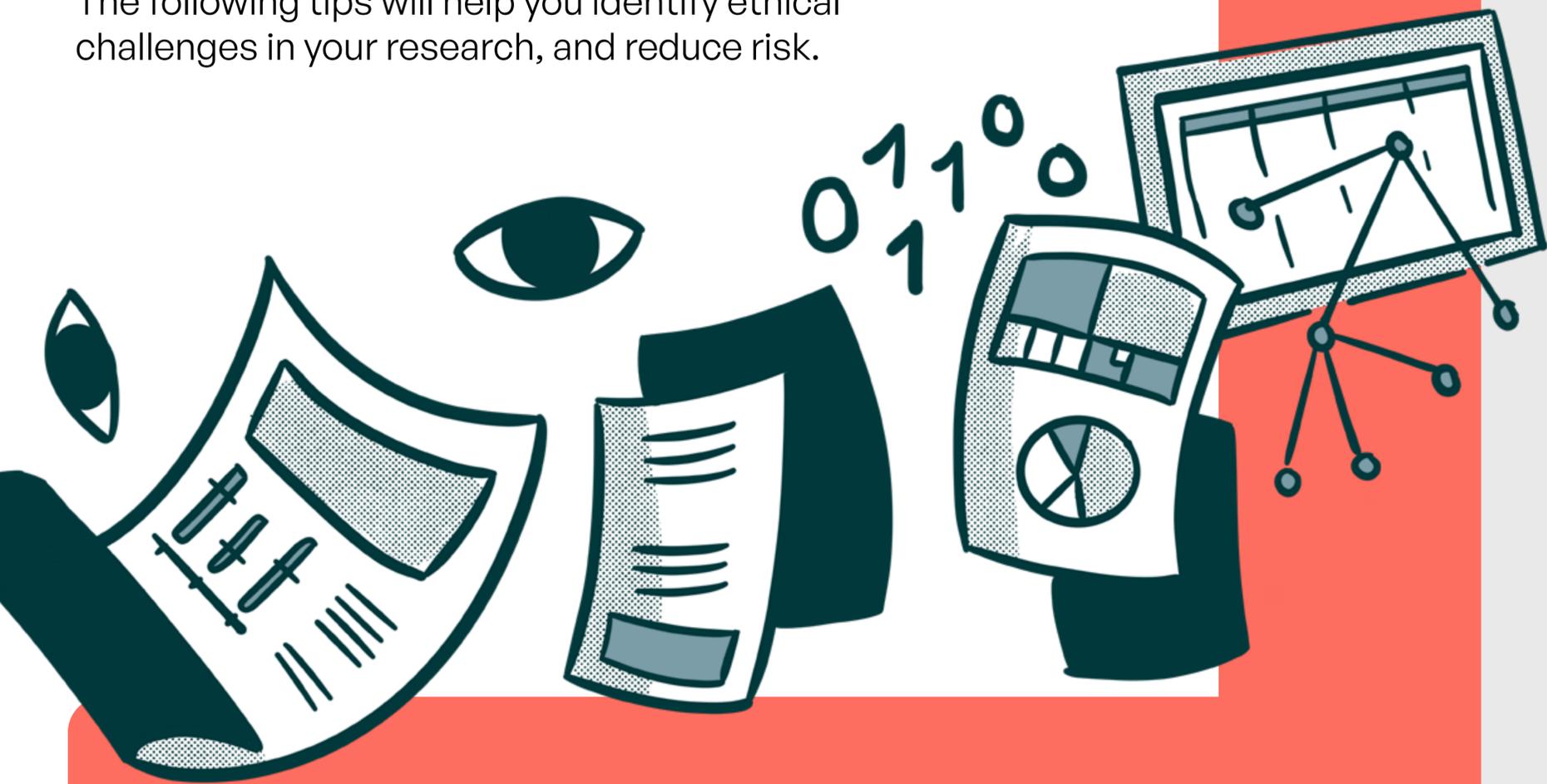
We employed a wide understanding of the term ethics to not artificially limit the scope of what CSS researchers viewed to be part of ethics in CSS research. Our goal was to develop best practices that were comprehensive and relevant to the full spectrum of ethical considerations in CSS research. To achieve this, we focused in particular to the nuances of ethical concerns. By using this approach, we were able to create a resource that is both grounded in empirical evidence and relevant to the needs of researchers in the field. We believe that this resource will be a valuable tool for promoting ethical research practices in CSS and supporting the work of ethical review boards in the region.



To carry out your research ethically, the research interest must outweigh risks of processing personal data. Personal data refers to an identified or identifiable individual.

But how do you know whether this is the case?

The following tips will help you identify ethical challenges in your research, and reduce risk.



BE AS CLEAR ABOUT THE RESEARCH QUESTION AND THE DATA NEEDED AS POSSIBLE

To enable weighting the research interest against the risk of processing personal data the research question should be known in as much detail and as early as possible in the research process.

If you expect additional questions to emerge throughout your research (e.g., in addition to studying hate speech patterns on a social media platform in Austria, you may also want to look at differences between genders, or rural and urban populations), think about what these might be from the start. It can save you headaches later on.

Rule of Thumb: The less granular the data, the better. The more granular the data, the more likely the risk of re-identification of involved persons or potential future misuse by other parties.



CONSULT WITH DOMAIN EXPERTS REGULARLY



Some technical properties that influence the results are only known to experts from the field (for example, mobility data from mobile phones may be recorded in less granularity in one state than in another, because utility poles there are located further apart from each other).

When unsure about interpreting any data consult a domain-expert. Discuss results with domain experts and see which ones surprised you but may not surprise them.

WHEN UNSURE CONSULT AN ETHICS REVIEW COMMITTEE

When unsure about whether your research entails questions of ethical relevance, consult with your research institution's Research Ethics Committee, Institutional Review Board, or research ethics experts.

If your institution does not have any of these, see if any of the involved researchers are affiliated with an institution that has institutional research ethics support and ask them to consult with them.





DISCUSS WITH PEERS

While your research is developing, discuss thoughts, concerns, and findings about the process with trusted peers.

If possible, set up regular discussion and reflection groups. You may learn from the experiences of your peers, and vice versa.



PREVENT MISINTERPRETATION OF RESULTS

Minimise the risk that your results will be misinterpreted. Put yourself in the shoes of an audience with little prior knowledge of the topic and think what they may get wrong.

Consider also how your findings could be deliberately misconstrued by actors who could benefit from them being misunderstood. For instance, presenting the findings of a study on Islamic extremism without mentioning other (e.g. right wing) extremism may give actors with bad intentions grounds to claim that extremism is always Islamic.

Pre-empt potential misinterpretation of findings by providing an in-depth discussion and contextualisation of your data and findings in your publications. You may also find it helpful to add information on what your findings do not mean.

Ask reviewers and colleagues for possible misinterpretations of your findings that you may have missed because you know the data and methods so well.



DO NOT ASSUME THAT PSEUDONYMISATION AND ANONYMISATION SOLVE ALL PROBLEMS

Removing identifiers (e.g. via pseudonymisation and anonymisation) are not infallible solutions.

In computational social science the meaning of these terms is shifting. There are many examples in which researchers thought they had anonymised the data, yet other researchers were able to identify research participants by linking large datasets.



VALIDATE YOUR RESULTS USING OTHER METHODS AND OTHER RELEVANT RESEARCH IF POSSIBLE

Discuss your findings with domain experts and compare your results to similar studies that employed different methods.



The methodology for collecting the best practices in research ethics in computational social science (CSS) was carefully designed to capture the perspectives of scholars in Austria, Germany, and Switzerland. The approach consisted of conducting 15 in-depth interviews, each lasting around an hour, with experienced CSS researchers. The interviewees were selected based on their expertise in the field and their demonstrated commitment to ethical research practices. To distil the most prominent ethical concerns and experiences of CSS researchers, we employed Kathy Charmaz's grounded theory approach. This method helped us identify key questions and themes that emerged from the interviews, which we then used to guide the development of the best practices. By using this approach, we ensured that the best practices were grounded in the real-world experiences and concerns of CSS researchers.

Illustrated by

MANUEL RECKER
Graphic Recording & Illustration