

## **An Ethical Overview**

The ethical principles of social research are anchored in the fundamental human rights, which are broadly formulated in the UN Declaration of Human Rights. Additional policies and declarations that codify principles of research ethics and the ethical treatment of research participants include the Nuremberg Code, the Helsinki Declaration, the Belmont Report, and the Menlo Report (Salganik 2017; European Commission 2018). These codes and addendums originate mostly in the biomedical field, though they encompass the central principles applied to all human research, which have led some academics to call for a Hippocratic oath for data scientists to safeguard against powerful new technologies under development in laboratories and tech firms (Rotblat 1999; Sample 2019). This discussion is nothing new however, as a tentative reformulation of the Hippocratic oath was introduced by Karl Popper (1969), wherein he stressed the importance of professional responsibility, a critical mind, and an overriding loyalty towards the betterment of mankind.

Matthew Salganik (2017) offer four principles deduced from the Belmont and Menlo Report that should guide the ethical deliberations of the researcher: 1) the respect for persons, that is individuals should be treated as autonomous and if circumstances require it individuals should be entitled to additional protections. 2) Beneficence stresses the importance of doing no harm and to maximize the possible benefits and minimizing any potential harms. 3) The principle of justice touches upon the importance of the distribution of burdens and benefits of the social scientist's research. This principle stress that is should not be a single stratum of society that bears the costs of the research while another stratum benefits. 4) The fourth and final principle is the respect for law and public interest, according to Salganik, the principle consists of two distinct elements, that is compliance to relevant laws and legal contracts and transparency-based accountability. It is worth noting that Popper's tentative Hippocratic oath mirrors the first three principles put forth by Salganik, stressing the importance of the ethical conduct of the researcher.

The need for rigid ethical standards within computational social science was made apparent by the Cambridge Analytica Scandal that broke on the 17th of March 2018. Where Steve Bannon could reveal that between 2013 and 2015 Cambridge Analytica had exploited a loophole in Facebook's API which allowed the company to harvest profile data from 87 million Facebook users, without the user's permission and use the harvested data to construct a massive targeted marketing database based on the user's likes and interests (Vox.com 2018). Other examples of misuse of data acquired from Facebook include the Harvard-run experiment, where students' data was used to create new knowledge about how social networks form and how these networks and their actors' behavior co-evolve, and the emotional contagion experiment from 2012, where approximately 700,000 users were involved in a research experiment to examine the extent to which a person's emotions are affected by the emotions of the people they interact with (Salgani 2017).

From this it should be evident that clear ethical guidelines are required in order to protect the user's privacy from tech-savvy companies. To this end, the European Parliament introduced the General Data Protection Regulation (GDPR). With its seven overarching principles the GDPR seeks to formalize the procedures involved in the data processing and storing of sensitive and private information (CRDC 2018). These principles include and expand upon the principles

found in the Belmont and Menlo Report. The most important consideration, however, must be that even a dataset comprising tens of thousands of observations involve human beings who must be protected from adverse side-effects of the social research. There is considerable evidence that point to the fact that even in anonymized data sets it can be possible to backtrack an individual's identity. A researcher must therefore be mindful of the mosaic effect if the dataset combines large amount of data from various sources (European Commission 2018).