



Ethics in Computational Social Science Research

Shuning Lu (Ph.D.)
Assistant Professor
Department of Communication
North Dakota State University

Email: shuning.lu@ndsu.edu
Twitter: [@shuning_lu](https://twitter.com/shuning_lu)



Why do we care about ethics in CSS?

Increasing power of researchers

Inconsistent rules, laws, and norms



Basic approaches to ethics

Rule-based (social science): Institutional Review Board

Ad-hoc (data science): little systematic experience



Principle-based approach

Respect for persons (informed consent)

Beneficence (do not harm, maximize possible benefits, minimize harms)

Justice (distribution of burdens and access to benefits)

Respect for law and public interest (i.e., comply with laws, be transparent)

Two frameworks of ethics

- Consequentialism: achieve ethical ends
- Deontology: use ethical means

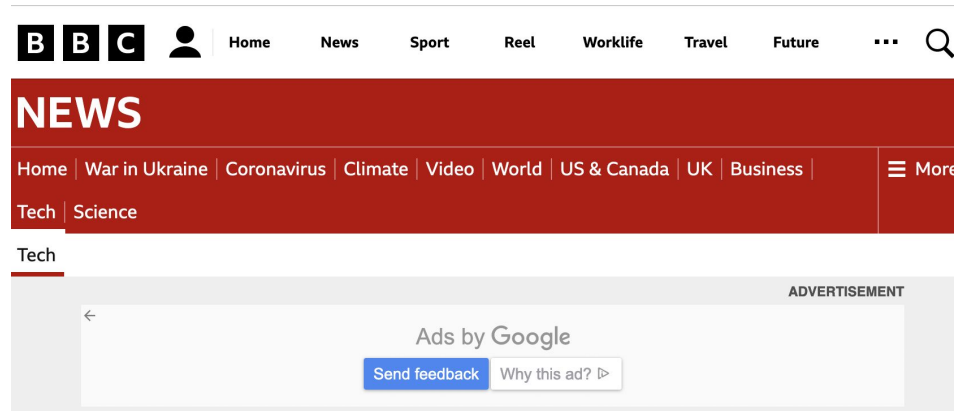
2014 Facebook Experiment

Informed consent?

Beneficence, potential harms vs. benefits?

Justice, distribution of burdens & access to benefits?

Respect for law and public interest



Facebook emotion experiment sparks criticism

30 June 2014 | Comments



Areas of difficulties #1

Informed consent for everything?

Abstract

Existing research on the extensive Chinese censorship organization uses observational methods with well-known limitations. We conducted the first large-scale experimental study of censorship by creating accounts on numerous social media sites, randomly submitting different texts, and observing from a worldwide network of computers which texts were censored and which were not. We also supplemented interviews with confidential sources by creating our own social media site, contracting with Chinese firms to install the same censoring technologies as existing sites, and—with their software, documentation, and even customer support—reverse-engineering how it all works. Our results offer rigorous support for the recent hypothesis that criticisms of the state, its leaders, and their policies are published, whereas posts about real-world events with collective action potential are censored.

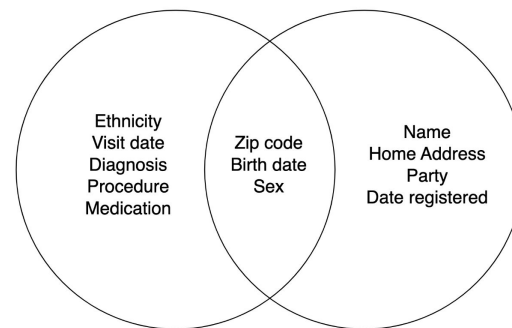
King, G., Pan, J., & Roberts, M. E. (2014). Reverse-engineering censorship in China: Randomized experimentation and participant observation. *Science*, 345(6199), 1251722.

Areas of difficulties #2

Informational risk (potential harm from disclosure of personal information)

Is removing PII enough?

Exposure after re-identification

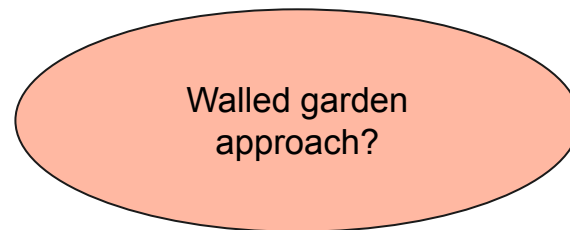
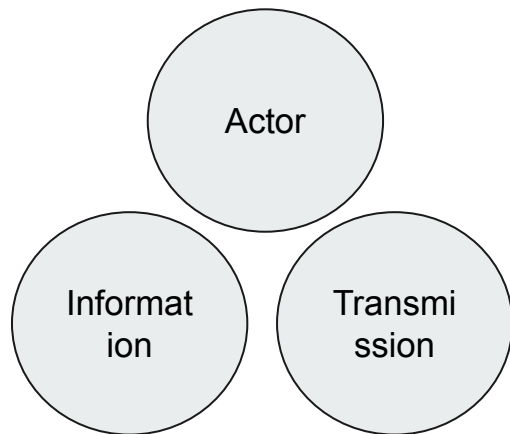


Re-identified medical data



Areas of difficulties #3

Privacy: from the public-private dichotomy to contextual integrity





Areas of difficulties #4

Decision-making under uncertainties

Unknown risks: unintended secondary uses



Practical tips

The IRB is a floor, not ceiling

Put yourself in everyone else's shoes (how do you want to be researched)

Think of ethics as continuous, not discrete



Small group activity Day 1

Activity

- Split into small groups.
- **Select one person** designated to take notes and report group discussion of case study
 - Read a [case study](#) of researchers facing a decision about research ethics.
 - Discuss and answer the four questions embedded in the case study.
 - Discuss and answer a fifth question: “Would you personally use the hacked data in this situation?”
- **Select a second person** designated to take notes and report discussion on second topic
 - Answer: “What other pressing ethical issues does your discipline or computational social science, in general, currently face?” You can discuss general issues and/or specific examples that have occurred in your field. Discuss as a group how to confront these issues.
- Come back together as a large group and discuss the case and questions.

Rough schedule

- 15 minutes: Introduce activity and form groups of 3 people
- 60 minutes: read and discuss case study in small groups; answer the questions
- 25 minutes: guided discussion of case study among the full group