Homework #3 (CMSC 396H, Spring 2017)

Due September 20, 9:00AM

1 Overview

The goal of this assignment is to investigate the ***ethics*** involved in CS research, while also gaining additional exposure into the kinds of research questions CS asks. We'll do this by comparing last week's paper ("All Your Contacts Are Belong to Us...") to the following paper:

> "Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests", S. Burnett, N. Feamster. In *ACM SIGCOMM*, 2015.

In class on September 20, we will have an interactive discussion about these papers, wherein we will discuss and debate the merits of the papers, with a particular emphasis towards to the ethical concerns these papers raise.

2 Ethics background

A major part of research is performing experiments, and often these experiments involve "human subjects." Some examples are relatively straightforward: human participants in a drug trial, or psychological experiments like the Stanford Prison Experiment. In the Stanford Prison Experiment, volunteers were randomly categorized into prisoners and prison guards, and they simulated running a prison on campus at Stanford. The goals of the research were to study how conflict arises among prisoners and prison guards. In a surprisingly short period of time, the "prison guards" began physically and psychologically torturing the "prisoners"—and even though they were allowed to opt out of the experiment at any time, almost all of the "prisoners" accepted the punishment. Eventually, an external party raised objections to those running the experiment, and they shut it down in less than a week.

Due to events such as these, the scientific community has adopted a set of guidelines and standards for how to go about running experiments that involve "human subjects." These are encompassed in the **Institutional Review Board (IRB) Process**. More information about the IRB process in general and in particular at UMD can be found here:

> http://www.umresearch.umd.edu/RCO/New/IRBProcess.html

And the federal guidelines for research involving human subjects can be found here:

> http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html

The IRB process concerns "human subjects" research, and seeks to ***protect the rights and welfare of "human subjects" participating in research studies.*** I keep putting "human subjects" in quotes because (a) it is an official term, and (b) particularly in CS, it is important to take a step back and consider what constitutes a human subject.

When does an experiment involve human subjects? For the experiments mentioned above, it is clear that they directly involve human subjects. And there are others where it is clear that no human subject is involved, such as when running a local simulation of protein folding on a computer. But then there are other experiments that involve humans in a somewhat indirect way—such as the Vanish paper, which ran experiments that interacted with software other humans were running, storing files on their computers. Does that involve human subjects?

What distinguishes these two classes of experiments? How do we know when there are human subjects involved? The HHS defines a human subject as follows:

> *Human subject* means a living individual about whom an investigator (whether professional or student) conducting research obtains
>
> 1. Data through intervention or interaction with the individual, or
> 2. Identifiable private information.

But what do "intervention" and "interaction" mean in CS, when we can interact with someone remotely, or cause software running on their machines to perform certain acts?

Informed consent When an experiment does involve a human subject, ***written informed consent*** must be obtained from all participants when possible (otherwise, the researchers must provide proper justification and obtain a waiver). Participants must be able to ask questions before, during, and after the experiment, and it must be the case that participation is always ***voluntary***.

Here, too, questions arise: when is it possible to obtain written consent and to field questions from participants in a CS research study, particularly when our experiments may include collecting data from hundreds of millions of users? What technical solutions are there to inform so many users, obtain consent, and field questions? Under what conditions should these rules be relaxed?

Risk versus benefit Ultimately, any experiment involving human subjects carries *some* risk with it. The federal guidelines for ethical research defines the notion of "minimal risk":

> *Minimal risk* means that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.

However, were we to require that *all* research impose no more than minimal risk, there are many worthwhile experiments that would not be able to take place: experimental drugs, new workout or training regimens, or even psychological analyses. So then where is the line?

The IRB process seeks to find a balance between the risk imposed on the human subject participants and the potential reward to society from the research. What is this balance? Whose decision is it? How do you quantify the risks from a CS research experiment, and the potential benefits? What role does the research community play?

Writing Task: Ethical review. For both papers under discussion ("All Your Contacts..." and "Encore..."), discuss the ethical aspects of the paper, touching on the following points (1–3 sentences each is fine, though if you want to expand on these points, of course feel free):

- Which experiments in this work involve human subjects? Why or why not? Did the paper acknowledge this, and if so, do you agree with their assessment?
- Should this work have sought out informed consent? Why or why not? If they sought it out, do you think they did so in an effective manner? How would you have gone about obtaining informed consent from the participants?
- What are the potential risks and benefits of the study? Did the paper describe them clearly? In your opinion, do the probable benefits outweigh the potential risks?
- Finally, include some concluding thoughts: Do you think that the paper (or *any* paper, for that matter) should be rejected based on ethical grounds alone (irrespective of the overall quality of the paper)? What role do you believe the Program Committee (PC) has in making such decisions?

3 Submitting

You may simply post your write-ups as *private* Piazza posts to me. These are due by 9AM the morning of our next class (Wednesday September 20).