# DRAFT – Ethics – DRAFT

August 18, 2016

## Contents

# 1 Ethics

## 1.1 Introduction

The previous chapters have shown that the digital age creates new opportunities for collecting and analyzing social data. The digital age has also created new ethical challenges. The goal of this chapter is to give you the tools that you need to handle these ethical challenges.

Their is currently uncertainty and disagreement about the appropriate conduct of some digital age social research. This uncertainty has led to two related problems, one of which has received much more attention than the other. On the one hand, some researchers have been accused of violating people's privacy or enrolling participants in unethical experiments. These cases—which I'll describe in this chapter—have been the subject of extensive debate and discussion. On the other hand, the ethical uncertainty has also had a chilling effect, preventing ethical and important research from happening; a fact that I think is much less appreciated. For example, during the 2014 Ebola outbreak, public health officials wanted information about the mobility of the people in the most heavily infected countries in order to help control the outbreak. Mobile phone companies had detailed call records that could have provided some of this information. Yet, ethical and legal concerns bogged down researchers' attempts to analyze the data (Wesolowski et al. 2014). If we can develop ethical norms and standards that are shared by both researchers and the public—and I think we can do this—then we can harness the capabilities of the digital age in ways that are responsible and beneficial to society.

There are important differences between how social scientists and data scientists approach research ethics. For social scientists, thinking about ethics is dominated by Institutional Review Boards (IRBs) and the regulations that they are tasked with enforcing. After all, the only way that most empirical social scientists experience ethical debate is through the bureaucratic process of IRB review. Data scientists, on the other hand, have little systematic experience with research ethics because it is not commonly discussed in computer science and engineering. Neither of these approaches—the *rules-based approach* of social scientists or the *ad-hoc approach* of data scientists—is well suited for social research in the digital age. Instead, I believe that we as as a community will make progress if we adopt a *principles-based approach*. That is, researchers should evaluate their research through existing rules—which I will take as given and assume should be followed—*and* through more general ethical principles. This *principles-based approach* ensures that researchers can make reasonable decisions about research for which rules have not yet been written and that we can communicate our reasoning with other researchers and the public.

The principles-based approach that I am advocating is not new; it draws on decades of previous thinking. As you will see, in some cases the principles-based approach leads to clear, actionable solutions. And, when it does not lead to such solutions, it clarifies the trade-offs involved, which is critical for striking an appropriate balance and being able to explain your reasoning to other researchers and the public. Further, as you will see, taking a principles-based approach does not require an inordinate amount of time. Once you learn the basic principles, you can use them to quickly and efficiently reason about a wide range of problems. Finally, the principles-based approach is sufficiently general that I expect that it will be helpful no matter where your research takes place or where you work (e.g., university, government, NGO, or company).

This chapter has been designed to help a well-meaning individual researcher. How should you think about the ethics of your own work? What can you do to make your own work more ethical? In Section 1.2, I'll describe three digital age research projects that have generated ethical debate. Then, in Section 1.3, I'll abstract from those specific examples to describe what I think is the fundamental reason for ethical uncertainty: rapidly increasing power for researchers to observe and experiment on people without their consent or even awareness. These capabilities are changing faster than our norms, rules, and laws. Next, in Section 1.4, I'll describe four existing principles that can guide your thinking: Respect for Persons, Beneficence, Justice, and Respect for Law and Public Interest. Then, in Section 1.5, I'll summarize two broad ethical frameworks—consequentalism and deontology—that can help you reason one of the deepest challenges that you might face: when is it appropriate for you to take ethically questionable means in order to achieve an ethically appropriate end. These principles and ethical frameworks will enable you to move beyond focusing on what is permitted by existing regulations and increase your ability to communicate your reasoning with other researchers and the public (Figure 1). With that background, in

Section 1.6, I will discuss four areas that are particularly challenging for digital age social researchers: informed consent (Section 1.6.1), understanding and managing information risk (Section 1.6.2), privacy (Section 1.6.3), and making ethical decisions in the face of uncertainty (Section 1.6.4). Finally, in Section 1.7, I'll conclude with three practical tips for working in an area with unsettled ethics. In the Historical Appendix, I'll describe the evolution of the current system of research ethics oversight in the United States including the Tuskegee Syphilis Study, the Belmont Report, the Common Rule, and the Menlo Report.

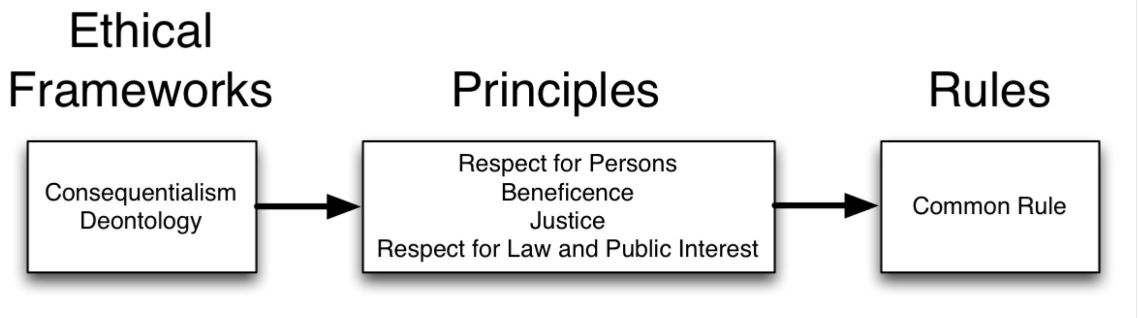


Figure 1: The rules governing research are derived from principles which in turn are derived from ethical frameworks. A main argument of this chapter is that researchers should evaluate their research through existing rules—which I will take as given and assume should be followed—*and* through more general ethical principles. The Common Rule is the set of regulations currently governing most federally-funded research in the United States (for more information, see the Historical Appendix). The four principles come from two blue-ribbon panels that have sought to provide ethical guidance to researchers: The Belmont Report and the Menlo Report (for more information, see the Historical Appendix). Finally, consequentialism and deontology are ethical frameworks that have been developed by philosophers for hundreds of years. A quick and crude way to distinguish the two frameworks is that consequentialists focus on ends and deontologists focus on means.

## 1.2 Three examples

**Digital age social research will involve studies where reasonable, well-meaning people will disagree about ethics.**

In order to keep the discussion of research ethics concrete, I'll start with three examples of digital age studies that have generated ethical controversy. I've selected these particular studies for two reasons. First, there are no easy answers about any of them. That is, I think that reasonable, well-meaning people will disagree about whether these studies should have happened and what changes might improve them. This lack of easy answers is characteristic of some studies today, and I expect that it will become more common in the future. Second, these three studies embody many of the principles, frameworks, and areas of tension that will follow later in the chapter.

### 1.2.1 Emotional Contagion

> **700,000 Facebook users were put into an experiment that may have altered their emotions. The participants did not give consent and the study was not subject to third-party ethical oversight.**

For one week in January of 2012, approximately 700,000 Facebook users were placed in an experiment to study emotional contagion, the extent to which a person's emotions are impacted by the emotions of the people they interact with. I've discussed this experiment in Chapter 4, but I'll review it again now. Participants in the Emotional Contagion experiment were put into four groups: a "negativity reduced" group, for whom posts with negative words (e.g., sad) were randomly blocked from appearing in the News Feed; a "positivity reduced" group for whom posts with positive words (e.g., happy) were randomly blocked; and two control groups. In the control for the "negativity reduced" group, posts were randomly blocked at the same rate as the "negativity reduced" group but without regard to the emotional content. The control group for the "positivity reduced" group was constructed in a parallel fashion. The researchers found that people in the positivity-reduced condition used slightly fewer positive words and slightly more negative words, relative to the control condition. Likewise, they found that people in the negativity-reduced condition used slightly more positive words and slightly fewer negative words. Thus, the researchers found evidence of emotional contagion (Kramer, Guillory, and Hancock 2014); for a more complete discussion of the design and results of the experiment see Chapter 4.

Just days after this paper was published in *Proceedings of the National Academy of Sciences*, there was an enormous outcry from both researchers and the press. Outrage around the paper focused on two main points: 1) participants did not provide any consent beyond the standard Facebook terms-of-service and 2) the study had not undergone third-party ethical review (Grimmelmann 2015). The ethical questions raised in this debate caused the journal to quickly publish a rare "editorial expression of concern" about the ethics and ethical review process for the research (Verma 2014). In subsequent years, this experiment has continued to be a source of intense debate and disagreement, and the criticism of this experiment may have had the unintended effect of driving this kind of research into the shadows (Meyer 2014). That is, some have argued that companies have not stopped running these kinds of experiments, they have merely stopped talking about them in public. This debate may have also lead to the creation of an ethical review process for research at Facebook (Hernandez and Seetharaman 2016; Jackman and Kanerva 2016).

### 1.2.2 Taste, Ties, and Time

> **Researchers scraped student data from Facebook, merged it with university records, used this merged data for research, and then shared it with other researchers.**

Beginning in 2006, each year a team of professors and research assistants scraped the Facebook profiles of all members of the Class of 2009 at a "diverse private college in the Northeastern U.S." This longitudinal data from Facebook on friendships and cultural tastes was merged with data the college had about the students' residential dorms and academic majors. This merged data represented a valuable resource for researchers, and it was used to create new knowledge about topics such as how social networks form (Wimmer and Lewis 2010) and how social networks and behavior co-evolve (Lewis, Gonzalez, and Kaufman 2012). In addition to using the data for their own work, the Taste, Ties, and Time research team made the data available to other researchers, after taking some steps to protect the students' privacy and in line with the wishes of the National Science Foundation (which funded the study) (Lewis et al. 2008).

Unfortunately, just days after the data were made available, other researchers deduced that the school in question was Harvard College (Zimmer 2010). The Taste, Ties, and Time researchers were accused of a "failure to adhere to ethical research standards" (Zimmer 2010) in part because the students had not provided informed consent (all procedures were reviewed and approved by Harvard's IRB and Facebook). In addition to criticism from academics, newspaper articles appeared with headlines such as "Harvard Researchers Accused of Breaching Students' Privacy" (Parry 2011). Ultimately, the dataset was removed from the Internet, and now it cannot be used by other researchers.

### 1.2.3 Encore

> **Researchers caused people's computers to secretly attempt to visit websites that were thought to be blocked by repressive governments.**

In March 2014, researchers launched Encore, a system to provide real-time and global measurements of Internet censorship. To understand how it worked, let's think about it in the context of your personal webpage (if you don't have one, you can imagine your friend's). One way to think about your webpage is as a computer program written in the html language. When a user visits your website, her computer downloads your html program and then renders it on her screen. Thus, your webpage is a program that is able to induce other people's computers to follow certain sets of instructions. Therefore, the researchers, Sam Burnett and Nick Feamster, who were at Georgia Tech, encouraged website owners to install a small code snippit into their webpages:

```
<iframe src="//encore.noise.gatech.edu/task.html"
        width="0" height="0"
        style="display: none"></iframe>
```

If you visit a webpage with this code snippit in it, here's what will happen. While your web browser was rending the webpage, the code snippit will cause your computer to try to contact a website that the researchers were monitoring. For example, it could be the website of a banned political party or persecuted religious group. Then, your computer will report back to the researchers about whether it was able to contact the potentially blocked website (Figure 2). Further, all of this would be invisible to you unless they checked the html source file of your webpage. Such invisible third-party page requests are actually quite common on the web (Narayanan and Zevenbergen 2015), but they rarely involve explicit attempts to measure censorship.

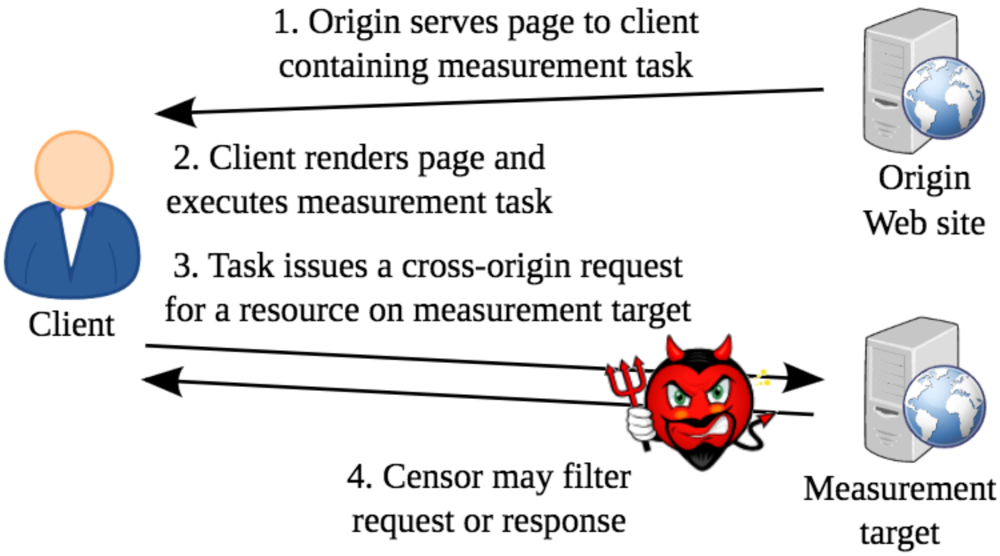


Figure 2: Schematic of the research design of Encore. The origin website sends you a webpage written in html with a small code snippet embedded in it (step 1). Your computer renders the webpage, which triggers the measurement task (step 2). Your computer attempts to access a measurement target, which could be the website of a banned political group (step 3). A censor, such as a government, may then block your access to the measurement target (step 4). Finally, your computer reports the results of this request to the researchers (not shown in the figure). Figure from Burnett and Feamster (2015).

This approach has some very attractive technical properties for measuring censorship. If enough websites add this code snippit, then the researchers can have a real-time, global-scale measure of which websites are censored by which countries. Before launching the project, the researchers conferred with the IRB at Georgia Tech, and the IRB declined to review the project because it was not "human subjects research" under the Common Rule (the Common Rule is the set of regulations governing most federally-funded research in the US; for more information, see the Historical Appendix at the end of this chapter).

Soon after Encore was launched, however, the researchers were contacted by Ben Zevenbergen, then a graduate student, who raised questions about the ethics of the project. In particular, there was a concern that people in certain countries could be exposed to risk if their computer attempted to visit certain sensitive websites, and these people who were being exposed to risk did not consent to participate in the study. Based on these conversations, the Encore team modified the project to only attempt to measure the censorship of Facebook, Twitter, and YouTube because third-party attempts to access these sites are common during normal web browsing (e.g., every webpage with a Facebook Like Button triggers a third-party request to Facebook).

After collecting data using this modified design, a paper describing the methodology and some results was submitted to SIGCOMM, a prestigious computer science conference. The program committee appreciated the technical contribution of the paper, but expressed concern about the lack of informed consent from participants. Ultimately, the program committee decided to publish the paper, but with a signing statement expressing ethical concerns (Burnett and Feamster 2015). Such a signing statement had never been used before at SIGCOMM, and this case has led to additional debate by computer scientists about the nature of ethics in their research (Narayanan and Zevenbergen 2015).

## 1.3 Digital is different

> **Social research in the digital age has different characteristics and therefore raises different ethical questions.**

Most social research in the analog age struck an appropriate ethical balance. For example, in a review of lab experiments that collectively involved more than 100,000 people, Plott (2013) found only one adverse event, a student who became upset because of losing money in an economic game. As the previous three digital age examples illustrate, however, researchers now face ethical challenges that are different from those in the past. Generalizing from these three studies, I think that the main problem that well-meaning researchers face is that capabilities are changing faster than rules, laws, and norms. More specifically, researchers—often in collaboration with companies and governments—have more power over participants than in the past. By power, I mean simply the ability to do things to people without their consent or even awareness. The things I'm talking about could be either observing their behavior or enrolling them in experiments. As the power of researchers to observe and perturb is increasing, there is not an equivalent increase in clarity about how that power should be used. In fact, researchers must decide how to exercise their power based on inconsistent and overlapping rules, laws, and norms. To be clear, this does not mean that most digital age research is unethical. In fact, given this situation, I think that researchers have shown remarkably good judgment. The combination of powerful capabilities and vague guidelines, however, puts well-meaning researchers in a difficult situation.

Although you personally might not feel especially powerful in terms of your ability to do things to people, increasingly researchers—often in collaboration with companies and governments—have the ability to observe and perturb people without their consent or awareness. For example, imagine following a person around and recording everything that they do. This would include tracking things such as where they go, what they buy, who they talk to, and what they read. Monitoring people like this in the analog age used to be the stuff of governments with enormous budgets. Now, all of this information is routinely and automatically recorded about millions and soon to be billions of people. Further, because all of this information is stored digitally, it is easy to copy, search, transmit, merge, and store. In other words, what is routinely done today would shock and amaze Cold War spy agencies like the KGB, CIA, and Stasi. Further, much of this behavioral tracking is taking place without the full understanding of those who are being surveilled.

A vivid metaphor that partially captures this situation of mass surveillance is the *panopticon*. First proposed in late 18th century by Jeremy Bentham as an architecture for prisons, the panopticon is the physical manifestation of surveillance (Figure 3). The panopticon is a circular building with rooms oriented around a central watchtower. Whoever occupies this watchtower can observe the behavior of all the people in the rooms. And, critically, the people in the rooms cannot observe the person in the watchtower. The person in the watchtower is thus an *unseen seer* (Foucault 1995).

In fact, digital surveillance is even more extreme than a person in a watchtower because it can produce a complete digital record of behavior that can be stored forever (Mayer-Schönberger 2009). While there is not yet a full recording of all human behavior merged into a master database, things are moving in that direction. And, that movement will most likely continue as long as the capabilities of sensors continue to increase, the cost of storage continues to decrease, and more of our lives become computer-mediated.

To many social researchers this master database might initially sound exciting, and it could certainly be used for a lot of important research. Legal scholars, however, have given a different name to this master database: *the database of ruin* (Ohm 2010). The creation of even an incomplete master database could have a chilling effect on social and political life if people become unwilling to read certain materials or discuss certain topics (Schauer 1978; Penney 2016). There is also a risk that the master database, while created for one purpose—say targeting ads—might one day be used for a different purpose, a situation called *secondary-use*. A horrific example of unanticipated secondary-use happened during the Second World War when government census data—the master database of that time—was used to facilitate the genocide that was taking place against Jews, Roma, and others (Table 1) (Seltzer and Anderson 2008). The statisticians who collected the data during peaceful times almost certainly had good intentions. But, when the world changed—when the Nazis came to power in Germany and neighboring countries—this data enabled a secondary-use was never intended. Once a master database exists, it is hard to anticipate who may gain access to it and how it will be used.

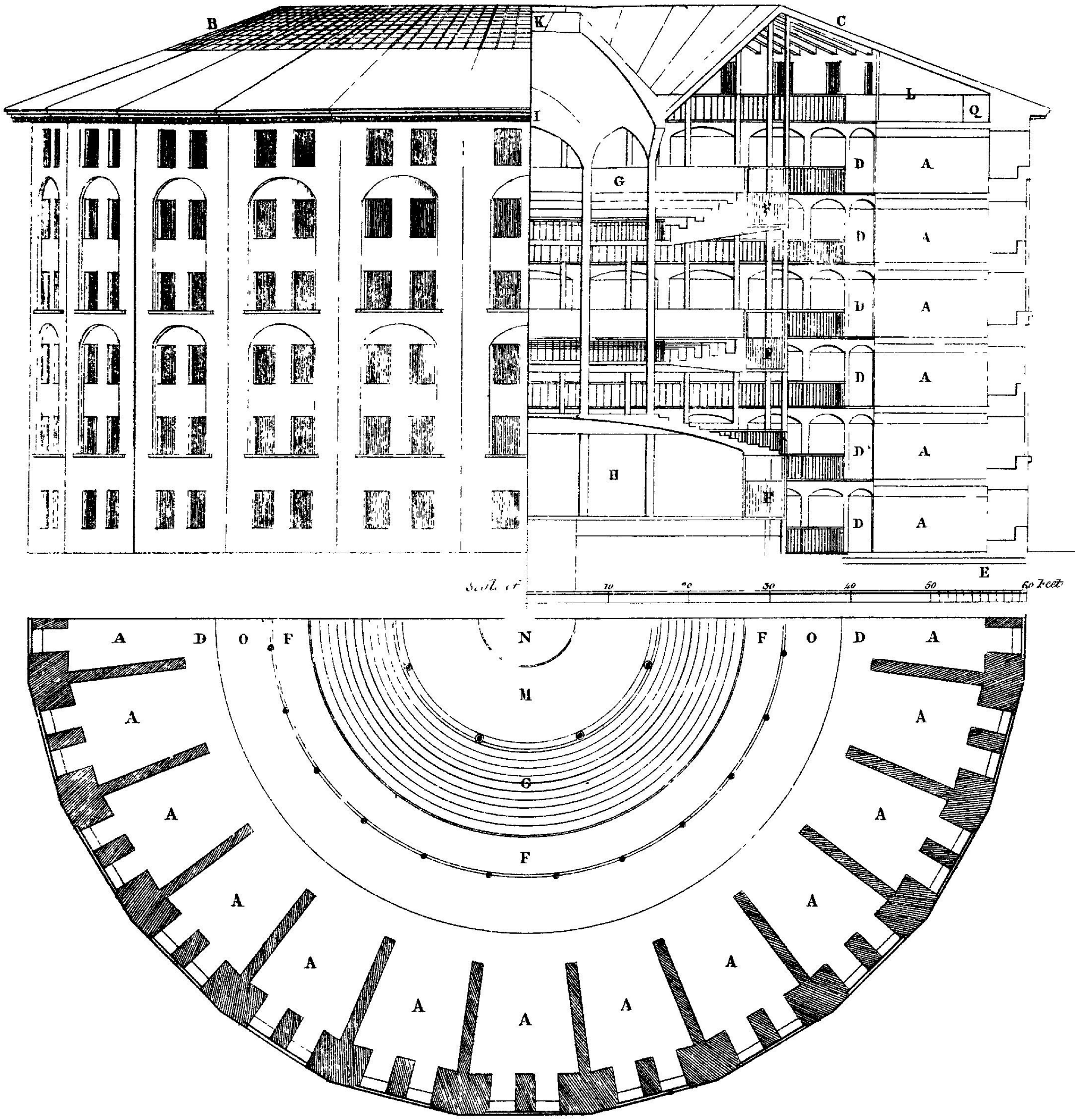


Figure 3: Design from the panopticon prison, first proposed by Jeremy Bentham. In the center, there is an unseen seer who can observe the behavior of everyone and cannot be observed. Drawing by Willey Reveley, 1791. Source: Wikimedia Commons.

Table 1: Cases where population data systems have been involved or potentially involved in human rights abuses. This table was original compiled by Seltzer and Anderson (2008), and I have included a subset of its columns. See Seltzer and Anderson (2008) for more information about each case and inclusion criteria. Some, but not all, of these cases involved unanticipated secondary use.

| Place | Time | Targeted individuals or groups | Data system | Human rights violation or presumed state intention |
|---|---|---|---|---|
| Australia | 19th & early 20th century | Aborigines | Population registration | Forced migration, elements of genocide |
| China | 1966-76 | Bad-class origin during cultural revolution | Population registration | Forced migration, instigated mob violence |
| France | 1940-44 | Jews | Population registration, special censuses | Forced migration, genocide |
| Germany | 1933-45 | Jews, Roma, and others | Numerous | Forced migration, genocide |
| Hungary | 1945-46 | German nationals and those reporting German mother tongue | 1941 Population Census | Forced migration |
| Netherlands | 1940-44 | Jews and Roma | Population registration systems | Forced migration, genocide |
| Norway | 1845-1930 | Samis and Kvens | Population censuses | Ethnic cleansing |
| Norway | 1942-44 | Jews | Special census & proposed population register | Genocide |
| Poland | 1939-43 | Jews | Primarily special censuses | Genocide |
| Romania | 1941-43 | Jews and Roma | 1941 Population Census | Forced migration, genocide |
| Rwanda | 1994 | Tutsi | Population registration | Genocide |

| Place | Time | Targeted individuals or groups | Data system | Human rights violation or presumed state intention |
|---|---|---|---|---|
| South Africa | 1950-93 | African and “Colored” popualtions | 1951 Population Census &population registration | Apartheid, voter disenfranchisement |
| United States | 19th century | Native Americans | Special censuses, population registers | Forced migration |
| United States | 1917 | Suspected draft law violators | 1910 Census | Investigation & prosecution of those avoiding registration |
| United States | 1941-45 | Japanese Americans | 1940 Census | Forced migration & internment |
| United States | 2001-08 | Suspected terrorists | NCES surveys & administrative data | Investigation & prosecution of domestic & international terrorists |
| United States | 2003 | Arab-Americans | 2000 Census | Unknown |
| USSR | 1919-39 | Minority populations | Various population censuses | Forced migration, punishment of other serious crimes |

Ordinary social researchers are very, very far from anything like creating chilling effects on society or participating in human right abuses through secondary-use. I’ve chosen to discuss these topics, however, because I think they will help social researchers understand the lens through which some people will see their work. Let’s return to the Taste, Ties, and Time project, for example. By merging together complete and granular data from Facebook with complete and granular data from Harvard, the researchers created an amazingly rich view of the social and cultural life of the students (Lewis et al. 2008). To many social researchers this seems like the master database, which could be used for good. But, to some others, it looks like the beginning of the database of ruin that was created without the consent of the participants. The Taste, Ties, and Time project began in 2006, and the information that researchers had was not particularly private. But, if you look forward a bit you can imagine that these issues are likely to get more complex. What kind of digital mosaic will researchers be able to construct about students in 2026 or 2046?

In addition to this mass surveillance, researchers—again in collaboration with companies and governments—can increasingly systematically intervene in people's lives in order to create randomized controlled experiments. For example, in Emotional Contagion, the researchers enrolled 700,000 people in an experiment without their consent or awareness. And, as I described in Chapter 5 (Running experiments), this kind of secret conscription of participants into experiments is not uncommon. Further, it does not require the cooperation of large companies. As I described in Chapter 5, researchers can increasingly design and build digital experiments with zero variable costs, a cost structure that enables extremely large experiments. Like the ability to observe, the ability to systematically perturb will likely continue to grow.

In the face of this increased power, researchers face *inconsistent and overlapping rules, laws, and norms.* One source of this inconsistency is that the capabilities of the digital age are changing more quickly than rules, laws, and norms. For example, the Common Rule (the set of regulations governing most government funded research in the United States) has changed little since 1981. An effort to modernize the Common Rule began in 2011 but was not complete as of the summer of 2016. A second source of inconsistency is that norms around abstract concepts like privacy are still being actively debated by researchers, policy makers, and activist. If specialists in these areas cannot reach uniform consensus, we should not expect that empirical researchers or participants will reach consensus either. A final source of inconsistency is that digital age research is increasingly mixed into other contexts, which leads to potentially overlapping norms and rules. For example, Emotional Contagion was a collaboration between a data scientist at Facebook and a professor and graduate student at Cornell. At Facebook running large experiments is routine as long as they comply with Facebook's terms of service, and at that time, there was no third-party review of experiments. At Cornell the norms and rules are quite different; virtually all experiments must be reviewed by the Cornell IRB. So, which set of rules should govern Emotional Contagion—Facebook's or Cornell's? When there are inconsistent and overlapping rules, laws, and norms even well-meaning researchers might have trouble doing the right thing. In fact, because of the inconsistency, there might not even be a single right thing.

Overall, these two features—increasing power and lack of agreement about how that power should be used—mean that researchers working in the digital age are going to face ethical challenges for the foreseeable future. Fortunately, researchers facing these challenges do not need to start from scratch. Instead, researchers can draw wisdom from previously developed ethical principles and frameworks, the topics of the next two sections.

## 1.4 Four principles

> **Four principles that can guide researchers facing ethical uncertainty are: Respect for Persons, Beneficence, Justice, and Respect for Law and Public Interest.**

The ethical challenges that researchers face in the digital age are somewhat different than those in the past. However, researchers can address these challenges by building on earlier ethical thinking. In particular, I believe that the principles expressed in two reports—The Belmont Report (Belmont Report 1979) and The Menlo Report (Dittrich, Kenneally, and others 2011)—can help researchers reason about the ethical challenges that they face. As I describe in more detail in the Historical Appendix, both of these reports were the results of multi-year deliberations by panels of experts with many opportunities for input from a variety of stakeholders.

First, in 1974, in response to ethical failures by researchers, such as the notorious Tuskegee Syphilis Study (see Historical Appendix), the US Congress created a national commission to write ethical guidelines for research involving human subjects. After four years of meeting at the Belmont Conference Center, the group produced the *Belmont Report*, a slender but powerful document. The Belmont Report is the intellectual basis for the *Common Rule*, the set of regulations governing human subjects research that Institutional Review Boards (IRBs) are tasked with enforcing (Porter and Koski 2008).

Then, in 2010, in response to the ethical failures of computer security researchers and the difficulty of applying the ideas in the Belmont Report to digital age research, the US Government—specifically the Department of Homeland Security—created a blue-ribbon commission to write a guiding ethical framework for research involving information and communication technologies (ICT). The results of this effort was the *Menlo Report* (Dittrich, Kenneally, and others 2011).

Together the Belmont Report and the Menlo Report offer four principles that can guide ethical deliberations by researchers: *Respect for Persons*, *Beneficence*, *Justice*, and *Respect for Law and Public Interest.* Applying these four principles in practice is not always straightforward, and it can require difficult balancing. The principles, however, help clarify trade-offs, suggest changes to research designs, and enable researchers to explain their reasoning to each other and the general public.

### 1.4.1 Respect for Persons

> **Respect for Persons is about treating people as autonomous and honoring their wishes.**

The Belmont Report argues that the principle of Respect for Persons consists of two distinct parts: (1) individuals should be treated as autonomous and (2) individuals with diminished autonomy should be entitled to additional protections. Autonomy roughly corresponds to letting people control their own lives. In other words, Respect for Persons suggests that researchers should not do stuff to people without their consent. Critically, this holds even if the researcher thinks that the thing that is happening is harmless or even beneficial. Respect for Persons leads to the idea that participants—not researchers—get to decide.

In practice, the principle of Respect for Persons has been interpreted to mean that researchers should, if possible, receive informed consent from participants. The basic idea with informed consent is that participants should be presented with relevant information in a comprehensible format and then should voluntarily agree to participate. Each of these terms has itself been the subject of substantial additional debate and scholarship (Manson and O'Neill 2007), and I'll devote an entire section later in this chapter to informed consent.

Applying the principle of Respect for Persons to the three examples from the beginning of the chapter highlights areas of concern with each of them. In each case, researchers did things to participants—used their data (Taste, Ties, or Time), used their computer to perform a measurement task (Encore), or enrolled them in an experiment (Emotional Contagion)—without their consent or awareness. The violation of the principle of Respect for Persons does not automatically make these studies ethically impermissible; Respect for Persons is one of four principles. But, thinking about Respect for Persons does suggest some ways that the studies could be improved ethically. For example, researchers could have gotten some form of consent from participants before the study began or after it ended; I'll return to these options when I discuss informed consent in more detail below. Finally, research ethicists emphasize that concerns about violating people's autonomy arise even in the case of completely benign studies. Concerns about harms and risks naturally enter ethical consideration, but they are generally addressed under the principle of Beneficence, the principle that I address next.

### 1.4.2 Beneficence

> **Beneficence is about understanding and improving the risk/benefit profile of your study, and then deciding if it strikes the right balance.**

The Belmont Report argues that the principle of Beneficence is an obligation that researchers have to participants, and that it involves two parts: (1) do not harm and (2) maximize possible benefits and minimize possible harms. The Belmont Report traces the idea of "do not harm" to the Hippocratic tradition in medical ethics, and it can be expressed in a strong form where researchers "should not injure one person regardless of the benefits that might come to others" (Belmont Report 1979). However, the Belmont Report also acknowledges that learning what is beneficial may involve exposing some people to risk. Therefore, the imperative of doing no harm can be in conflict with the imperative to learn, leading researchers to make occasionally difficult decisions about "when it is justifiable to seek certain benefits despite the risks involved, and when the benefits should be foregone because of the risks." (Belmont Report 1979)

In practice, the principle of Beneficence has been interpreted to mean that researchers should undertake two separate processes: a risk/benefit analysis and then a decision about whether the risks and benefits strike an appropriate ethical balance. This first process is largely a technical matter requiring substantive expertise, and the second is largely an ethical matter where substantive expertise may be less valuable or even detrimental.

A risk/benefit analysis involves both understanding *and* improving the risks and benefits of a study. Analysis of risk should include two elements: the probability of adverse events and the severity of those events. During this stage, for example, a researcher could adjust the study design to reduce the probability of an adverse event (e.g., screen out participants who are vulnerable) or reduce the severity of an adverse event if it occurs (e.g., make counseling available to participants who request it). Further, during this process researchers need to keep in mind the impact of their work not just on participants, but also on non-participants and social systems. For example, consider the experiment by Restivo and van de Rijt (2012) on the effect of awards on Wikipedia editors (discussed in Chapter 4). In this experiment, the researchers gave awards to some editors that they considered deserving and then tracked their contributions to Wikipedia compared to a control group of equally deserving editors to whom the researchers did not give an award. In this particular study, the number of awards they gave was small, but if the researchers had flooded Wikipedia with awards it could have disrupted the community of editors without harming any of them individually. In other words, when doing risk/benefit analysis you should think about the impacts of your work not just on participants but on the world more broadly.

Next, once the risks have been minimized and the benefits maximized, researchers should assess whether the study strikes a favorable balance. Ethicists do not recommend a simple summation of costs and benefits. In particular, some risks render the research impermissible no matter the benefits (e.g., the Tuskegee Syphilis Study described in the Historical Appendix). Unlike the risk/benefit analysis, which is largely technical, this second step is deeply ethical and may in fact be enriched by people who do not have specific subject-area expertise. In fact, because outsiders often notice different things from insiders, IRBs in the US are required to have at least one non-researcher. In my experience serving on an IRB, these outsiders can be helpful for preventing group-think. So if you are having trouble deciding whether your research project strikes an appropriate risk/benefit analysis don’t just ask your colleagues, try asking some non-researchers; their answers might surprise you.

Applying the principle of Beneficence to the three examples highlights the fact that there is often substantial uncertainty about risks before a study begins. For example, the researchers did not know the probability or magnitude of the adverse events that could be caused by their studies. This uncertainty is actually quite common in digital age research, and later in this chapter, I'll devote an entire section to the challenge of making decisions in the face of uncertainty (Section 1.6.4). However, the principle of Beneficence does suggest some changes that might be made to these studies to improve their risk/benefit balance. For example, in Emotional Contagion, the researchers could have attempted to screen out people under 18 years old and people who might be especially likely to react badly to the treatment. They could have also tried to minimize the number of participants by using efficient statistical methods (as described in detail in Chapter 4). Further, they could have attempted to monitor participants and offer assistance to anyone that appeared to have been harmed. In Taste, Ties, and Time, the researchers could have put extra safeguards in place when they released the data (although their procedures were approved by Harvard's IRB which suggests that they were consistent with common practice at that time); I'll offer some more specific suggestions about data release later in the chapter when I describe informational risk (Section 1.6.2). Finally, in Encore, the researchers could have attempted to minimize the number of risky requests that are created in order to achieve the measurement goals of the project, and they could have excluded participants that are most in danger from repressive governments. Each of these possible changes would introduce trade-offs into the design of these projects, and my goal is not to suggest that these researchers should have made these changes. Rather, my goal is to show the kinds of changes that the principle of Beneficence can suggest.

Finally, although the digital age has generally made the weighing of risks and benefits more complex, it has actually made it easier for researchers to increase the benefits of their work. In particular, the tools of the digital age greatly facilitate open and reproducible research, where researchers make their research data and code available to other researchers and make their papers available to the public by publishing open access. This change to open and reproducible research, while by no means simple, offers a way for researchers to increase the benefits of their research without exposing participants to any additional risk (data sharing is an exception that will be discussed in detail in the section on informational risk (Section 1.6.2)).

### 1.4.3 Justice

> **Justice is about ensuring that the risks and benefits of research are distributed fairly.**

The Belmont Report argues that the principle of Justice addresses the distribution of the burdens and benefits of research. That is, it should not be the case that one group in society bears the costs of research while another group reaps its benefits. For example, in the 19th and early 20th century, the burdens of serving as research subjects in medical trials fell largely on the poor, and the benefits of improved medical care flowed primarily to the rich.

In practice, the principle of Justice was initially interpreted around the idea that vulnerable people should be protected from researchers. In other words, researchers should not be allowed to intentionally prey on the powerless. It is a troubling pattern that in the past, a large number of ethically problematic studies have involved extremely vulnerable participants including poorly educated and disenfranchised citizens (Jones 1993); prisoners (Spitz 2005); institutionalized, mentally disabled children (Robinson and Unruh 2008); and old and debilitated hospital patients (Arras 2008).

Around 1990, however, views of Justice began to swing from *protection* to *access* (Mastroianni and Kahn 2001). For example, activists argued that children, women, and ethnic minorities needed to be explicitly included in clinical trials so that these groups could benefit from the knowledge gained by these trials.

In addition to questions about protection and access, the principle of Justice is often interpreted to raise questions about appropriate compensation for participants—questions which are subject to intense debate in medical ethics (Dickert and Grady 2008).

Applying the principle of Justice to the three examples offers yet another way to evaluate them. Participants in none of the studies were compensated financially. Encore raises the most complex questions about the principle of Justice. While the principle of

Beneficence might suggest excluding participants from countries with repressive governments, the principle of Justice could argue against denying these people the possibility to participant in–and benefit from—accurate measurements of Internet censorship. The case of Taste, Ties, and Time also raises questions. In this case, one group of students bears the burdens of the research and society as a whole benefits. To be clear, however, this was not an especially vulnerable population. Finally, in Emotional Contagion the participants were a sample from the population most likely to benefit from the results of the research, a situation well aligned with the principle of Justice.

### 1.4.4 Respect for Law and Public Interest

> **Respect for Law and Public Interest extends the principle of Beneficence beyond specific research participants to include all relevant stakeholders.**

The fourth and final principle that can guide your thinking is Respect for Law and Public Interest. This principle comes from the Menlo Report, and therefore may be less well known to social researchers. The Menlo Report argues that the principle of Respect for Law and Public Interest is implicit in the principle of Beneficence, but the Menlo Report argues that it deserves explicit consideration. In my mind, the best way to think about this principle is that Beneficence tends to focus on participants and that Respect for Law and Public Interest explicitly encourages researchers to take a wider view and include law in their considerations. In analog age research—such as traditional surveys and lab experiments—researchers were unlikely to accidentally break the law. In online research, this is, unfortunately, much less true.

In the Menlo Report, Respect for Law and Public Interest has two distinct components: (1) *Compliance* and (2) *Transparency-based Accountability*. Compliance means that researchers attempt to identify and obey relevant laws, contracts, and terms of service. For example, compliance would mean that a researcher considering scraping the content of a website should read and consider the terms-of-service agreement of that website. There may, however, be situations where it is permissible to violate the terms of service. For example, at one time both Verizon and AT&T had terms of service that prevented customers from criticizing them (Vaccaro et al. 2015). Researchers should not be automatically bound by such terms-of-service agreements. Ideally, if researchers violate terms of service agreements, they should explain their decision openly (e.g., Soeller et al. (2016)). But, this openness may expose researchers to added legal risk. In the United States, for example, the Computer Fraud and Abuse Act makes it illegal to violate terms of service agreements (Sandvig and Karahalios 2016).

Further, transparency-based accountability means that researchers need to be clear about the goals, methods, and results at all stages of their research process and to take responsibility for their actions. Another way to think about this transparency-based accountability is that it is trying to prevent the research community from doing things in secret. This transparency-based accountability enables a broader role for the research community and the public in ethical debates, which is important for both ethical and practical reasons.

Applying the principle of Respect for Law and Public Interest to these three studies illustrates some of the complexity researchers face when it comes to law. For example, Grimmelmann (2015) has argued that Emotional Contagion may have been illegal under law in the State of Maryland. In particular, Maryland House Bill 917, passed in 2002, extends Common Rule protections to all research conducted in Maryland, independent of funding source (recall that many experts believe that Emotional Contagion was not subject to the Common Rule under Federal Law because it was conducted at Facebook, an institution that does not receive research funds from the US Government). However, some scholars believe that Maryland House Bill 917 is itself unconstitutional [Grimmelmann (2015); p. 237-238]. Practicing social researchers are not judges, and therefore are not equipped to understand and assess the constitutionality of the laws of all 50 US states. These complexities are compounded in international projects. Encore, for example, involved participants from 170 countries, which makes legal compliance incredibly difficult. In response to the ambiguous legal environment, researchers should be careful to undergo third-party ethical review of their work, as both a source of advice about legal requirements and as a personal protection in case their research is unintentionally illegal.

On the other hand, all three studies published their results in academic journals enabling transparency-based accountability. In fact, Emotional Contagion was published open access so the research community and the broader public were informed—after the fact—about the design and results of the research. One rule of thumb to assess transparency-based accountability is to ask yourself: would I be comfortable if my research procedures were written about on the front page of my home town newspaper? If the answer is no, that is a strong sign that your research design needs changes.

In conclusion, the Belmont Report and Menlo Report propose four principles that can be used to assess research: Respect for Persons, Beneficence, Justice, and Respect for Law and Public Interest. Applying these four principles in practice is not always straightforward, and it can require difficult balancing. For example, when deciding whether to debrief participants from Emotional Contagion, Respect for Persons might encourage debriefing whereas Beneficence might discourage debrief (if the debriefing would itself do harm). There is no automatic way to balance these competing principles, but at a minimum, the four principles help clarify trade-offs, suggest changes to research designs, and enable researchers to explain their reasoning with each other and the general public.

## 1.5 Two ethical frameworks

> **Most debates about research ethics reduce to disagreements between consequentialism and deontology.**

These four ethical principles are themselves largely derived from two more abstract ethical frameworks: *consequentialism* and *deontology.* Understanding these frameworks is helpful because it will help you identify and then reason about one of the most fundamental tensions in research ethics: when can you use potentially unethical means to achieve an ethical end.

Consequentialism, which has roots in the work of Jeremy Bentham and John Stuart Mill, focuses on taking actions that lead to better states in the world (Sinnott-Armstrong 2014). The principle of Beneficence, which focuses on balancing risk and benefits, is deeply rooted in consequentialist thinking. On the other hand, deontology, which has roots in the work of Immanuel Kant, focuses on ethical duties, independent of their consequences

(Alexander and Moore 2015). The principle of Respect for Persons, which focuses on the autonomy of participants, is deeply rooted in deontological thinking. A quick and crude way to distinguish the two frameworks is that consequentialists focus on *ends* and deontologists focus on *means*.

To see how these two frameworks can differ, consider informed consent. Both frameworks could be used to support informed consent but for different reasons. A consequentialist argument for informed consent is that it helps to prevent harm to participants by prohibiting research that does not properly balance risk and anticipated benefit. In other words, consequentialist thinking would support informed consent because it helps prevent bad outcomes for participants. However, a deontological argument for informed consent focuses on a researcher's duty to respect the autonomy of her participants. Given these approaches, a pure consequentialist might be willing to waive the requirement for informed consent in a setting where there was no risk, whereas a pure deontologist might not.

Both consequentialism and deontology offer important ethical insight, but each can be taken to absurd extremes. For consequentialism, one of these extreme cases could be called *Transplant*. Imagine a doctor who has five patients dying of organ failure and one healthy patient whose organs can save all five. Under certain conditions, a consequenalist doctor will be permitted—and even required—to kill the healthy patient to obtain his organs. This complete focus on ends, without regard to means, is flawed.

Likewise, deontology can also be taken to awkward extremes, such as in the case that could be called *Timebomb*. Imagine a police officer who has captured a terrorist who knows the location of a ticking timebomb that will kill millions of people. A deontological police officer would not lie in order to trick a terrorist into revealing the location of the bomb. This complete focus on means, without regards to ends, also is flawed.

In practice, most social researchers implicitly embrace a blend of these two ethical frameworks. Noticing this blending of ethical schools helps clarify why many ethical debates—which tend to be between those who are more consequentialist and those who are more deontological—don't make much progress. These debates rarely resolve because consequentialists offer arguments about ends, arguments that are not convincing to deontologists who are worried about means. Likewise, deontologists tend to offer arguments about means, which are not convincing to consequentialists who are focused on ends. Arguments between consequentialists and deontologists are like two ships passing in the night.

One solution to these debates would be for social researchers to develop a consistent, morally solid, and easy-to-apply blend of consequentialism and deontology. Unfortunately, that's unlikely to happen; philosophers have been working on these problems for a long time. Therefore, I think the only course of action is to acknowledge that we are working from inconsistent foundations and muddle forward.

## 1.6 Areas of difficulty

The four ethical principles—Respect for Persons; Beneficence; Justice; and Respect for Law and Public Interest—and the two ethical frameworks—consequential and deontology—should help you reason about any research ethics problems that you are facing. However, based on the characteristics of digital age research described earlier in this chapter and based on the ethical debates we have seen so far, I see four areas of particular difficulty: *informed consent*, *understanding and managing informational risk*, *privacy*, and *making decisions in the face of uncertainty*. In the next sections, I will describe these four issues in more detail and offer advice about how to handle them.

### 1.6.1 Informed consent

> **Researchers should, can, and do follow the rule: some form of consent for most research.**

Informed consent is a foundational idea—some might say a near obsession (Emanuel, Wendler, and Grady 2000; Manson and O'Neill 2007)—in research ethics. The simplest version of research ethics says: "informed consent for everything." This simple rule, however, is *not* consistent with existing ethical principles, ethical regulation, or research practice. Instead, researchers should, can, and do follow a more complex rule: "some form of consent for most research."

First, in order to move beyond overly simplistic ideas about informed consent, I want to tell you more about field experiments to study discrimination (these were covered a bit in

Chapter 4 as well). In these studies, fake applicants who have different characteristics—such as a man or women—apply for different jobs. If one type of applicant gets hired at a higher rate, then researchers can conclude that there may be discrimination in the hiring process. For the purposes of this chapter, the most important thing about these experiments is that the participants in these experiments—the employers—never provide consent. In fact, these participants are actively deceived by the fake applications. Yet, field experiments to study discrimination have been performed in at least 117 studies in 17 countries (Riach and Rich 2002; Rich 2014).

Researchers who use field experiments to study discrimination have identified four features of these studies that, collectively, make them ethically permissible: 1) the limited harm to the employers, 2) the great social benefit of having reliable measure of discrimination, 3) the weakness of other methods of measuring discrimination, and 4) the fact that deception does not strongly violate the norms of that setting (Riach and Rich 2004). Each of these conditions is critical, and had any of them not been satisfied, the ethical case would be more challenging. Three of these features can be derived from the ethical principles in the Belmont Report: limited harm (Respect for Persons and Beneficence) and great benefit and weakness of other methods (Beneficence and Justice). The final feature, non-violation of contextual norms, can be derived from the Menlo Report's Respect for Law and Public Interest. In other words, employment applications are a setting where there is already some expectation of possible deception. Thus, these experiments do not pollute an already pristine ethical landscape.

In addition to this principles-based argument, dozens of IRBs have also concluded that the lack of consent in these studies is consistent with existing rules, in particular Common Rule §46.116, part (d). Finally, US courts have also supported the lack of consent and deception in field experiments to measure discrimination (No. 81-3029. United States Court of Appeals, Seventh Circuit). Thus, the use of field experiments without consent is consistent with existing ethical principles and existing rules (at least the rules in the US). This reasoning has been supported by the broad social research community, dozens of IRBs, and by the US Court of Appeals. Thus, we must reject the simple rule "informed consent for everything." This is not a rule that researchers follow, nor is it a rule that researchers should follow.

Moving beyond "informed consent for everything" leaves researchers with a difficult question: what forms of consent are needed for what kinds of research? Naturally, there has been substantial prior debate around this question, although most of it is in the context of medical research in the analog age. Summarizing that debate, Eyal (2012) writes:

> "The more risky the intervention, the more it is a high-impact or a definitive 'critical life choice', the more it is value-laden and controversial, the more private the area of the body that the intervention directly affects, the more conflicted and unsupervised the practitioner, the higher the need for robust informed consent. On other occasions, the need for very robust informed consent, and indeed, for consent of any form, is lesser. On those occasions, high costs may easily override that need." [internal citations excluded]

An important insight from this prior debate is that informed consent is not all or nothing; there are stronger and weaker forms of consent. In some situations, robust informed consent seems necessary, but, in other situations, weaker forms of consent may be appropriate. Next, I'll describe three situations where researchers will struggle to obtain informed consent from all affected parties and a few options in those cases.

First, sometimes asking participants to provide informed consent may increase the risks that they face. For example, in Encore, asking people living under repressive governments to provide consent to have their computer used for measurement of Internet censorship might place those who agree at increased risk. When consent leads to increased risk, researchers can ensure that information about what they are doing is public and that it is possible for participants to opt-out. Also, they could seek consent from groups that represent the participants (e.g., NGOs).

Second, sometimes having fully informed consent before the study begins could compromise the scientific value of the study. For example, in Emotional Contagion, if participants had known that researchers were doing an experiment about emotions, this might have changed their behavior. Withholding information from participants, and even deceiving them, is not uncommon in social research, especially in lab experiments in psychology. If informed consent is not possible before a study begins, researchers could (and usually do) *debrief* participants after the study is over. Debriefing general includes explaining what actually happened, remediating any harms, and obtaining consent after the fact. There is some debate, however, about whether debriefing in field experiments is appropriate, if the debriefing itself can cause harm to participants (Finn and Jakobsson 2007).

Third, sometimes it is logistically impractical to receive informed consent from everyone impacted by your study. For example, imagine a researcher who wishes to study the Bitcoin blockchain (Bitcoin is a crypto-currency and the blockchain is a record of all Bitcoin transactions (Narayanan et al. 2016)). Some people that use Bitcoin expect and desire anonymity, and some members of the Bitcoin community might object to certain forms of research on their community. Unfortunately, it is impossible to obtain consent from everyone who uses Bitcoin because many of these people are anonymous. In this case, the researcher could try to contact a sample of Bitcoin users and ask for their informed consent.

These three reasons why researchers might not be able to receive informed consent—increasing risk, compromising research goals, and logistical limitations—are not the only reasons that researchers struggle to obtain informed consent. And, the solutions that I've suggested—informing the public about your research, enabling an opt-out, seeking consent from third-parties, debriefing, and seeking consent from a sample of participants—may not be possible in all cases. Further, even if these alternatives are possible, they may not be sufficient for your study. What these examples do show, however, is that informed consent is not all or nothing, and that creative solutions can improve the ethical balance of studies that cannot receive full informed consent from all impacted parties.

To conclude, rather than "informed consent for everything" researchers should, can, and do follow a more complex rule: "some form of consent for most things." Expressed in terms of principles, informed consent is neither necessary nor sufficient for the principles of Respect for Persons [Humphreys (2015); p. 102]. Further, Respect for Persons is just one of the principles that needs to be balanced when considering research ethics; it should not automatically overwhelm Beneficence, Justice, and Respect for Law and Public Interest, a point made repeatedly by ethicists over the past 40 years [Gillon (2015); p. 112-113]. Expressed in terms of ethical frameworks, informed consent for everything is an overly deontological position that falls victim to situations such as *Timebomb* (see Section 1.5).

Finally, as a practical matter, if you are considering doing research without any kind of consent, then you should know that you are in a gray area. Be careful. Look back at the ethical argument that researchers have made in order to conduct experimental studies of discrimination without consent. Is your justification as strong? Because informed consent is central to many lay ethical theories, you should know that you will likely be called on to defend your decisions.

### 1.6.2 Understanding and managing informational risk

> **Information risk is the most common risk in social research; it has increased dramatically; and it is the hardest risk to understand.**

The second ethical challenge for social age digital research is *informational risk*, the potential for harm from the disclosure of information (Council 2014). Informational harms from the disclosure of personal information could be economic (e.g., losing a job), social (e.g., embarrassment), psychological (e.g., depression), or even criminal (e.g., arrest for illegal behavior). Unfortunately, the digital age increases information risk dramatically—there is just so much more information about our behavior. And, informational risk has proven very difficult to understand and manage compared to risks that were concerns in analog age social research, such as physical risk. To see how the digital age increases informational risk, consider the transition from paper to electronic medical records. Both types of records create risk, but the electronic records create much greater risks because at a massive scale they can be transmitted to an unauthorized party or merged with other records. Social researchers in the digital age have already run into trouble with informational risk, in part because they didn't fully understand how to quantify and manage it. So, I'm going to offer a helpful way to think about informational risk, and then I'm going to give you some advice for how to manage the informational risk in your research and in releasing data to other researchers.

One way that social researchers decrease informational risk is *"anonymization"* of data. "Anonymization" is the process of removing obvious personal identifiers such as name, address, and telephone number from the data. However, this approach is much less effective than many people realize, and it is, in fact, deeply and fundamentally limited. For that reason, whenever I describe "anonymization," I'll use quotation marks to remind you that this process creates the appearance of anonymity but not true anonymity.

A vivid example of the failure of "anonymization" comes from the late 1990s in Massachusetts (Sweeney 2002). The Group Insurance Commission (GIC) was a government agency responsible for purchasing health insurance for all state employees. Through this work, the GIC collected detailed health records about thousands of state employees. In an effort to spur research about ways to improve health, GIC decided to release these records to researchers. However, they did not share all of their data; rather, they "anonymized" it by removing information such as name and address. However, they left other information that they thought could be useful for researchers such as demographic information (zip code, birth date, ethnicity, and sex) and medical information (visit data, diagnosis, procedure) (Figure 4) (Ohm 2010). Unfortunately, this "anonymization" was not sufficient to protect the data.

To illustrate the shortcomings of the GIC "anonymization", Latanya Sweeney—then a graduate student at MIT—paid $20 to acquire the voting records from the city of Cambridge, the hometown of Massachusetts governor William Weld. These voting records included information such as name, address, zip code, birth date, and gender. The fact that the medical data file and the voter file shared fields—zip code, birth date, and sex—meant that Sweeney could link them. Sweeney knew that Weld's birthday was July 31, 1945, and the voting records included only six people in Cambridge with that birthday. Further, of those six people, only three were male. And, of those three men, only one shared Weld's zip code. Thus, the voting data showed that anyone in the medical data with Weld's combination of birth date, gender, and zip code was William Weld. In essence, these three pieces of information provided a *unique fingerprint* to him in the data. Using this fact, Sweeney was able to locate Weld's medical records, and to inform him of her feat, she mailed him a copy of his records (Ohm 2010).

Sweeney's work illustrates the basic structure of *de-anonymization attacks*—to adopt a term from the computer security community. In these attacks, two data sets, neither of

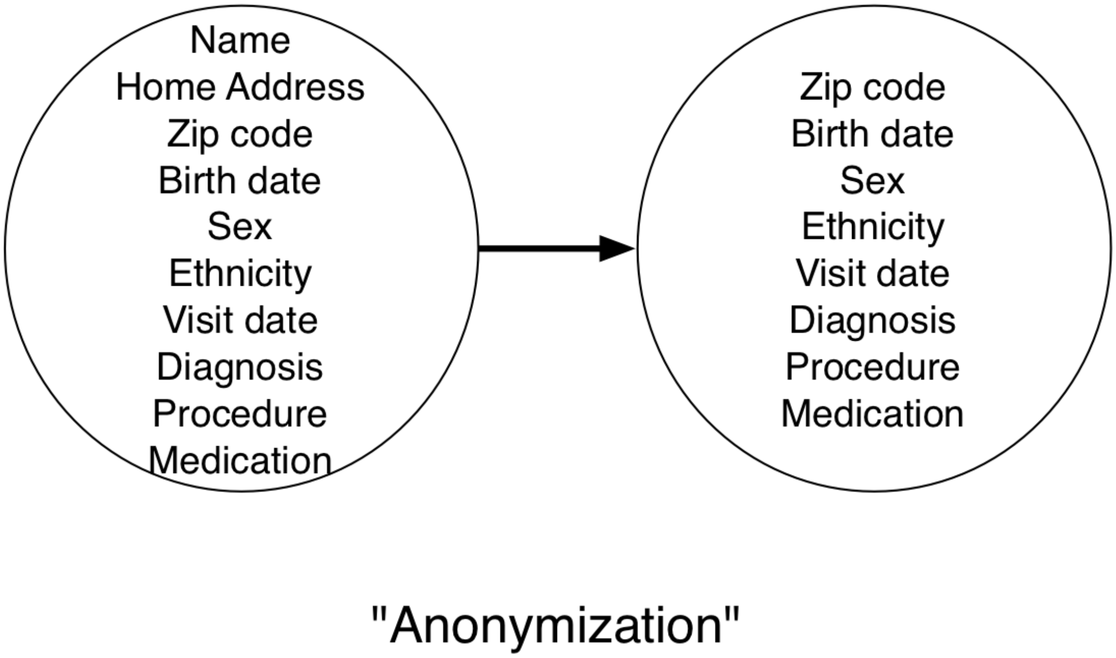


Figure 4: “Anonymization” is the process of removing obviously identifying information. For example, when releasing the medical insurance records of state employees the Massachusetts Group Insurance Commission (GIC) removed name and address from the files. I use quotes around the word “anonymization” because the process provides the appearance of anonymity, but not actual anonymity.

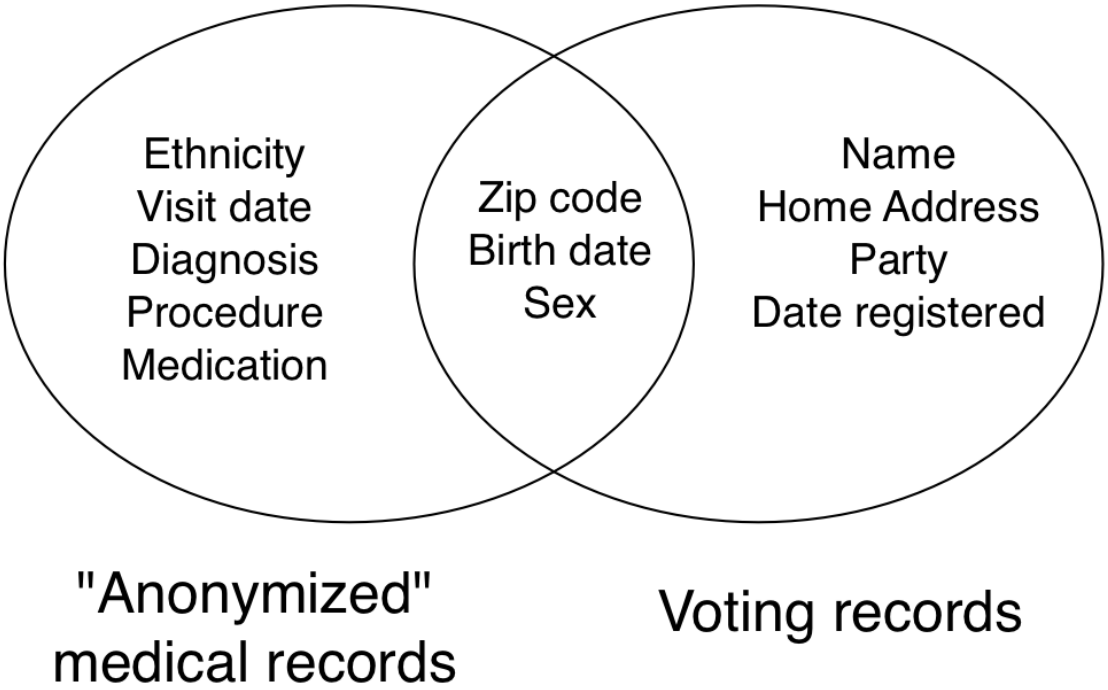


Figure 5: Re-idenification of “anonymized” data. Latanya Sweeney combined the “anonymized” health records with voting records in order to find the medical records of Governor William Weld (Sweeney 2002).

which by itself reveals sensitive information, are linked, and through this linkage, sensitive information is exposed. In some ways this process is similar to the way that baking soda and vinegar, two substances that are by themselves safe, can be combined to produce a nasty outcome.

In response to Sweeney's work, and other related work, researchers now generally remove much more information—all so called "Personally Identifying Information" (PII) (Narayanan and Shmatikov 2010)—during the process of "anonymization." Further, many researchers now realize that certain data—such as medical records, financial records, answers to survey questions about illegal behavior—is probably too sensitive to release even after "anonymization." However, more recent examples that I'll describe below indicate that social researchers need to change their thinking. As a first step, it is wise to assume that all data is *potentially* identifiable and all data is *potentially* sensitive. In other words, rather than thinking that informational risk applies to a small subset of projects, we should assume that it applies—to some degree—to all projects.

Both aspects of this re-orientation are illustrated by the Netflix Prize. As described in Chapter 5, Netflix released 100 million movie ratings provided by almost 500,000 members, and had an open call where people from all over the world submitted algorithms that could improve Netflix's ability to recommend movies. Before releasing the data, Netflix removed any obviously personally-identifying information, such as names. Netflix also went an extra step and introduced slight perturbations in some of the records (e.g., changing some ratings from 4 stars to 3 stars). Netflix soon discovered, however, that despite their efforts, the data were by no means anonymous.

Just two weeks after the data were released Narayanan and Shmatikov (2008) showed that it was possible to learn about specific people's movie preferences. The trick to their re-identification attack was similar to Sweeney's: merge together two information sources, one with potentially sensitive information and no obviously identifying information and one that contains the identity of people. Each of these data sources may be individually safe, but when they are combined the merged dataset can create informational risk. In the case of the Netflix data, here's how it could happen. Imagine that I choose to share my thoughts about action and comedy movies with my co-workers, but that I prefer not to share my opinion about religious and political movies. My co-workers could use the information that I've shared with them to find my records in the Netflix data; the information that I share could be a *unique fingerprint* just like William Weld's birth date, zip code, and sex. Then, if they find my unique fingerprint in the data, they could learn my ratings about all movies, including movies where I choose not to share. In addition to this kind of *targeted attack* focused on a single person, Narayanan and Shmatikov (2008) also showed that it was possible to do a *broad attack*—one involving many people—by merging the Netflix data with personal and movie rating data that some people have chosen to post on the Internet Movie Database (IMDb). Any information that is unique fingerprint to a specific person—even their set of movie ratings—can be used to identify them.

Even though the Netflix data can be re-identified in either a targeted or broad attack, it still might appear to be low risk. After all, movie ratings don't seem very sensitive. While that might be true in general, for some of the 500,000 people in the dataset, movie ratings might be quite sensitive. In fact, in response to the de-anonymization a closeted lesbian woman joined a class-action suit against Netflix. Here's how the problem was expressed in their lawsuit (Singel 2009):

> "[M]ovie and rating data contains information of a more highly personal and sensitive nature [sic]. The member's movie data exposes a Netflix member's personal interest and/or struggles with various highly personal issues, including sexuality, mental illness, recovery from alcoholism, and victimization from incest, physical abuse, domestic violence, adultery, and rape."

The de-anonymization of the Netflix Prize data illustrates both that all data is potentially identifiable and that all data is potentially sensitive. At this point, you might think that this only applies to data that that purports to be about people. Surprisingly, that is not the case. In response to a Freedom of Information Law request, the New York City Government released records of every taxi ride in New York in 2013, including the pickup and drop off times, locations, and fare amounts (recall from Chapter 2 that Farber (2015) used this data to test important theories in labor economics). Although this data about taxi trips might seem benign because it does not seem to be information about people, Anthony Tockar realized that this taxi dataset actually contained lots of potentially sensitive information about people. To illustrate, he looked at all trips starting at The Hustler Club—a large strip club in New York—between midnight and 6am and then found their drop-off locations. This search revealed—in essence—a list of addresses of some people who frequent The Hustler Club (Tockar 2014). It is hard to imagine that the city government had this in mind when it released the data. In fact, this same technique could be used to find the home addresses of people who visit any place in the city—a medical clinic, a government building, or a religious institution.

These two cases—the Netflix Prize and the New York City taxi data—show that relatively skilled people failed to correctly estimate the informational risk in the data that they released, and these cases are by no means unique (Barbaro and Zeller Jr 2006; Zimmer 2010; Narayanan, Huey, and Felten 2016). Further, in many of these cases, the problematic data is still freely available online, indicating the difficulty of ever undoing a data release. Collectively these examples—as well as research in computer science about privacy—leads to an important conclusion. Researchers should assume that all data is *potentially* identifiable and all data is *potentially* sensitive.

Unfortunately, there is no simple solution to the fact that all data is potentially identifiable and all data is potentially sensitive. However, one way to reduce information risk while you are working with data is to create and follow a *data protection plan*. This plan will decreases the chance that your data will leak and will decrease the harm if a leak somehow occurs. The specifics of data protection plans, such as which form of encryption to use, will change over time, but the UK Data Services helpfully organizes the elements of a data protection plan into 5 categories that they call the *5 safes*: safe projects, safe people, safe settings, safe data, and safe outputs (Table 2) (Desai, Ritchie, and Welpton 2016). None of the five safes individually provide perfect protection. But, together they form a powerful set of factors that can decrease informational risk.

Table 2: The 5 safes are principles for designing and executing a data protection plan (Desai, Ritchie, and Welpton 2016).

| Safe | Action |
|---|---|
| Safe projects | limits projects with data to those that are ethical |
| Safe people | access is restricted to people who can be trusted with data (e.g., people have undergone ethical training) |
| Safe data | data is de-identified and aggregated to the extent possible |
| Safe settings | data is stored in computers with appropriate physical (e.g., locked room) and software (e.g., password protection, encrypted) protections |
| Safe output | research output is reviewed to prevent accidentally privacy breaches |

In addition to protecting your data while you are using it, one step in the research process where informational risk is particularly salient is data sharing with other researchers. Data sharing among scientists is a core value of the scientific endeavor, and it greatly facilities the advancement of knowledge. Here's how the UK House of Commons described the importance of data sharing:

> "Access to data is fundamental if researchers are to reproduce, verify and build on results that are reported in the literature. The presumption must be that, unless there is a strong reason otherwise, data should be fully disclosed and made publicly available. In line with this principle, where possible, data associated with all publicly funded research should be made widely and freely available." (Molloy 2011)

Yet, by sharing your data with another researcher, you may be increasing informational risk to your participants. Thus, it may seem that researchers who wish to share their data—or are required to share their data—are facing a fundamental tension. On the one hand they have an ethical obligation to share their data with other scientists, especially if the original research is publicly funded. Yet, at the same time, researchers have an ethical obligation to minimize, as much as possible, the information risk to their participants.

Fortunately, this dilemma is not as severe as it appears. It is important to think of data sharing along a continuum from *no data sharing* to *release and forget*, where data is "anonymized" and posted for anyone to access (Figure 6). Both of these extreme positions have risks and benefits. That is, it is not automatically the most ethical thing to not share your data; such an approach eliminates many potential benefits to society. Returning to Taste, Ties, and Time, an example discussed earlier in the chapter, arguments against data release that focus only on possible harms and that ignore possible benefits are overly one-sided; I'll describe the problems with this one-sided, overly protective approach in more detail in below when I offer advice about making decisions in the face of uncertainty (Section 1.6.4).

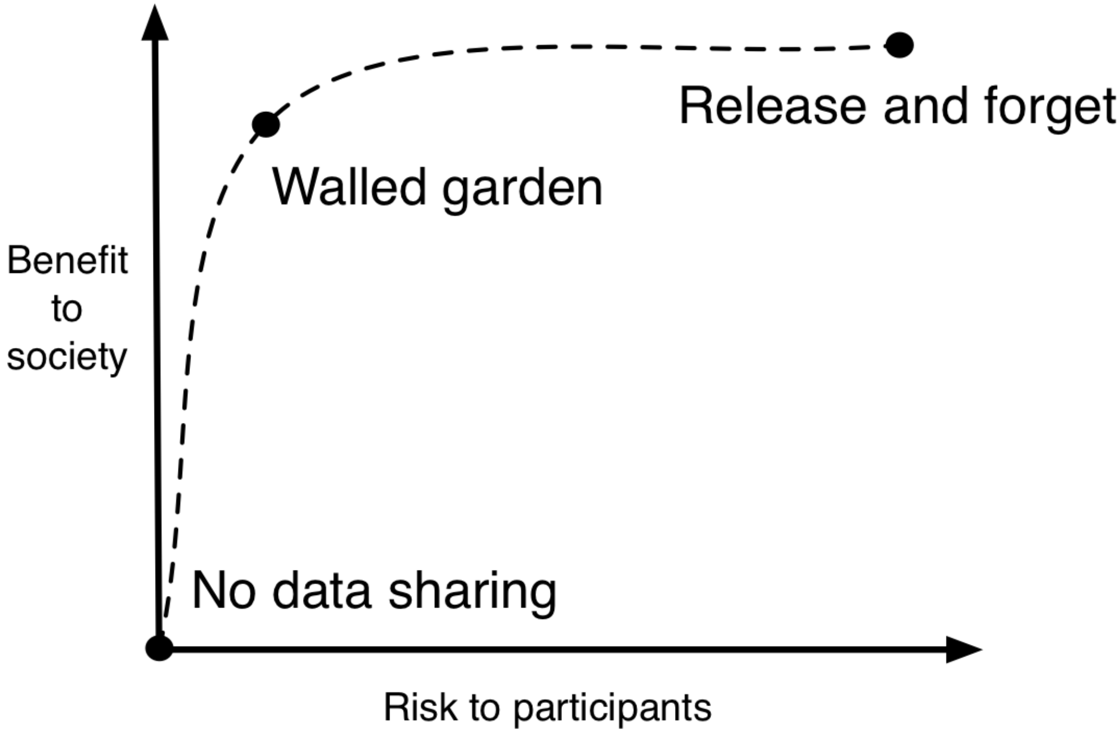


Figure 6: Data release strategies can fall along a continuum. Where you should be along this continuum depends on the specific details of your data. In this case, third party review may help you decide the appropriate balance of risk and benefit in your case.

Further, in between these two extreme cases is what I'll called a *walled garden* approach where data is shared with people who meet certain criteria and who agree to be bound by certain rules (e.g., oversight from an IRB and a data protection plans). This walled garden approach provides many of the benefits of release and forget with less risk. Of course, a walled garden approach creates many questions—who should have access, under what conditions, for how long, who should pay to maintain and police the walled garden etc.—but these are not insurmountable. In fact, there are already working walled gardens in place that researchers can use right now, such as the data archive of the Inter-university Consortium for Political and Social Research at the University of Michigan.

So, where should the data from your study be on the continuum of no sharing, walled garden, and release and forget? It depend on the details of your data; researchers must balance Respect for Persons, Beneficence, Justice, and Respect for Law and Public Interest. When assessing appropriate balance for other decisions researchers seek the advice and approval of IRBs, and data release can be just another part of that process. In other words, although some people think of data release as a hopeless ethical morass, we already have systems in place to help researchers balance these kind of ethical dilemmas.

One final way to think about data sharing is by analogy. Every year cars are responsible for thousands of deaths, but we do not attempt to ban driving. In fact, such a call to ban driving would be absurd because driving enables many wonderful things. Rather, society places restrictions on who can drive (e.g., need to be a certain age, need to have passed certain tests) and how they can drive (e.g., under the speed limit). Society also has people tasked with enforcing these rules (e.g., police), and we punish people who are caught violating them. This same kind of balanced thinking that society applies to regulating driving can also be applied to data sharing. That is, rather than making absolutist arguments for or against data sharing, I think the biggest benefits will come from figuring out how we can share more data more safely.

To conclude, informational risk has increased dramatically, and it is very hard to predict and quantify. Therefore, it is best to assume that all data is potentially identifiable and potentially sensitive. To decrease informational risk while doing research, researchers can create and follow a data protection plan. Further, informational risk does not prevent researchers from sharing data with other scientists.

### 1.6.3 Privacy

**Privacy is a right to the appropriate flow of information.**

A third area where researchers may struggle is *privacy*. As Lowrance (2012) put it quite succinctly: “privacy should be respected because people should be respected.” Privacy, however, is notoriously messy concept (Nissenbaum 2010, Ch. 4), and as such, it is difficult to use when trying to make specific decisions about research.

A common way to think about privacy is with a public/private dichotomy. By this way of thinking, if information is publicly accessible, then it can be used by researchers without concerns about violating people’s privacy. But this approach can run into problems. For example, in November 2007 Costas Panagopoulos sent everyone in three towns a letter about an upcoming in election. In two towns—Monticello, Iowa and Holland, Michigan—Panagopoulos promised/threatened to publish a list of people who had voted in the newspaper. In the other town—Ely, Iowa—Panagopoulos promised/threatened to publish a list of people who had not voted in the newspaper. These treatments were designed to induce pride and shame (Panagopoulos 2010) because these emotions had been found to impact turnout in earlier studies (Gerber, Green, and Larimer 2008). Information about who votes and who doesn’t is public in the United States; anyone can access it. So, one could argue that because this voting information is already public, there is no problem with the researcher publishing it in the newspaper. On the other hand, something about that argument feels wrong to many people.

As this example illustrates, the public/private dichotomy is too blunt (boyd and Crawford 2012; Markham and Buchanan 2012). A better way to think about privacy, one especially designed to handle issues raised by the digital age, is the idea of *contextual integrity* (Nissenbaum 2010). Rather than considering information public or private, contextual integrity focuses on the flows of information. For example, many people would be unbothered if their doctor shared their health records with another doctor but would be unhappy if their doctor sold this same information to a marketing company. Thus, according to Nissenbaum (2010), “a right to privacy is neither a right to secrecy or a right to control but a right to *appropriate flow* of personal information.”

The key concept underlying contextual integrity is *context-relative informational norms* (Nissenbaum 2010). These are norms that govern the flows of information in specific settings, and they are determined by three parameters:

- actors (subject, sender, recipient)
- attributes (types of information)
- transmission principles (constraints under which information flows)

Thus, when you as a researcher are deciding whether to use data without permission it is helpful to ask, "Does this use violate context-relative informational norms?" Returning to the case of Panagopoulos (2010), in this case, having an outside researcher publish lists of voters or non-voters in the newspaper seems likely to violate informational norms. In fact, Panagopoulos did not follow through on his promise/threat because local election officials traced the letters to him and persuaded him that it was not a good idea (Issenberg 2012, 307).

In other settings, however, thinking about context-relative informational norms requires a bit more consideration. For example, let's return to the possibility of using mobile phone call logs to track mobility during the Ebola outbreak in West Africa in 2014, a case that I discussed in the introduction to this chapter (Wesolowski et al. 2014). In this setting, we can imagine two different situations:

- Situation 1: sending complete call log data [attributes]; to governments of incomplete legitimacy [actors]; for any possible future use [transmission principles]
- Situation 2: sending partially anonymized records [attributes]; to respected university researchers [actors]; for use in response to the Ebola outbreak and subject to the oversight of university ethical boards [transmission principles]

Even though in both of these situations call data are flowing out of the company, the informational norms concerning these two situations are not the same because of differences between the actors, attributes, and transmission principles involved. Focusing on only one of these parameters can lead to overly simplistic decision-making. In fact, Nissenbaum (2015) emphasizes that none of these three parameters can be reduced to the others, nor can any one of them individually define informational norms. This three-dimensional nature of informational norms explains why past efforts—that have focused on either attributes or transmission principles—have been ineffective at capturing common-sense notions of privacy.

One challenge with using the idea of context-relative informational norms to guide decisions is that researchers might not know them ahead of time and they are very hard to measure (Acquisti, Brandimarte, and Loewenstein 2015). Further, even if some research would violate contextual-relative informational norms that does not automatically mean that the research should not happen. In fact, Chapter 8 of Nissenbaum (2010) is entirely about "Breaking Rules for Good." Despite these complications, context-relative informational norms is still a very useful way to reason about questions related to privacy.

Finally, privacy is an area where I've seen many misunderstandings between researchers who prioritize Respect for Persons and those who prioritize Beneficence. Imagine the case of a public health researcher who secretly watches people taking showers because understanding hygiene is key to preventing the spread of a novel infectious disease. Researchers focusing on Beneficence would focus on the benefits to society from this research and might even argue that there is no harm to participants if the researcher does her spying without detection. On the other hand, researchers who prioritize Respect for Persons would focus on the fact that the researcher is not treating people with respect and is in fact doing them harm by violating their privacy. Unfortunately, it is not easy to resolve the conflicting views of this situation (although the best solution in this case might just be to ask for consent).

In conclusion, when reasoning about privacy, it is helpful to move beyond the overly simplistic public/private dichotomy and to reason instead about context-relative informational norms, which are made of up three elements: actors (subject, sender, recipient), attributes (types of information), and transmission principles (constraints under which information flows) (Nissenbaum 2010). Some researchers evaluate privacy in terms of harm that could result from the violation of privacy, whereas other researchers view the violation of privacy as a harm in and of itself. Because notions of privacy in many digital systems are changing over time, vary from person to person, and vary from situation to situation (Acquisti, Brandimarte, and Loewenstein 2015), privacy is likely to be a source of difficult ethical decisions for researchers for some time.

### 1.6.4 Making decisions in the face of uncertainty

**Uncertainty need not lead to inaction.**

The fourth and final area where I expect researchers to struggle is making decisions in the face of uncertainty. That is, after all the philosophizing and balancing, research ethics involves making decisions about what to do and what not to do. Unfortunately, these decisions often must be made based on incomplete information. For example, when designing Encore, researchers might wish to know the probability that it will cause someone to be visited by the police. Or, when designing Emotional Contagion researchers might wish to know the probability that it could trigger depression in some participants. These probabilities are probably extremely low, but they are unknown before the research takes place. And, because neither project publicly tracked information about adverse events, these probabilities are not generally known even after the projects were completed.

Uncertainties are not unique to social research in the digital age. The Belmont Report, when describing the systematic assessment of risks and benefits, explicitly acknowledges these will be difficult to quantify exactly. These uncertainties, however, are more severe in the digital age, in part because we have less experience, and in part because of the characteristics of digital age social research.

Given these uncertainties some people seem to advocate for something like "better safe than sorry," which is a colloquial version of the *Precautionary Principle*. While this approach appears reasonable—perhaps even wise—it can actually cause harm; it is chilling to research; and it causes people think in the wrong way (Sunstein 2005). In order to understand the problems with the Precautionary Principle, let's consider Emotional Contagion. The experiment was planned to involve about 700,000 people, and there was certainly some chance that people in the experiment would suffer harm. But, there was also some chance the experiment could yield knowledge that would be beneficial to Facebook users and to society. Thus, while allowing the experiment is a risk (as has been amply discussed), preventing the experiment is also a risk because the experiment could have produced valuable knowledge. Of course, the choice is not between doing the experiment as it occurred and not doing the experiment; there are many possible modifications to the design that might have brought it into a different ethical balance. However, at some point, researchers will have the choice between doing a study and not doing a study, and there are risks in both action and inaction. It is inappropriate to focus only on the risks of action. Quite simply, there is no risk-free approach.

Moving beyond the Precautionary Principle, one important way to think about making decisions given uncertainty is the *minimal risk standard*. The minimal risk standard attempts to benchmark the risk of a particular study against the risks that participants undertake in their daily lives, such as playing sports and driving cars (Wendler et al. 2005). This approach is valuable because assessing whether something is minimal risk is easier than assessing the actual level of risk. For example, in Emotional Contagion, before the research started, the researchers could have compared the emotional content on naturally occurring News Feeds to the emotional content that participants would see in the experiment (Meyer 2015). If the News Feeds under the treatment were similar to those that naturally occur on Facebook, then the researchers could conclude that the experiment is minimal risk. And, they could make this decision *even if they don't know the absolute level of risk*. The same approach could be applied to Encore. Initially, Encore triggered requests to websites that were known to be sensitive, such as websites of banned political groups in countries with repressive governments. As such, it was not minimal risk for participants in certain countries. However, the revised version of Encore—which only triggered requests to Twitter, Facebook, and YouTube—is minimal request because requests to those sites are triggered during normal web browsing (Narayanan and Zevenbergen 2015).

A second important idea is when making decisions about studies with unknown risk is *power analysis*, which allows researchers to calculate an appropriate size for their study (Cohen 1988). That is, if your study might expose participants to risk—even minimal risk—then the principle of Beneficence suggests that you want to impose the smallest amount of risk needed to achieve your research goals. (Think back to the Reduce principle that I discussed in Chapter 4.) Even though some researchers have an obsession with making their studies as *big* as possible, research ethics suggests that we should make our studies as *small* as possible. Thus, even if you don't know the exact level of risk your study involves, a power analysis can help you ensure that it is as small as possible. Power analysis is not new, of course, but there is an important difference between the way that it was used in the analog age and how it should be used today. In the analog age, researchers generally did power analysis to make sure that their study was not too small (i.e., under-powered).

Now, however, researchers should do power analysis to make sure that their study is not too big (i.e., over-powered). If you do a power analysis and your study appears to require an enormous number of people, then that may be a sign that the effect you are studying is tiny. If so, you should ask whether this small effect is sufficiently important to impose a large number of people to risks of an unknown size. In many situations the answer is probably no (Prentice and Miller 1992).

The minimal risk standard and power analysis help you reason about and design studies, but they don't provide you any new information about how participants might feel about your study and what risks they might experience from participating in your study. Another way to deal with uncertainty is to collect additional information, which leads to ethical-response surveys and staged trials.

In *ethical-response surveys*, researchers present a brief description of a proposed research project and then ask two questions:

- (Q1) "If someone you cared about were a candidate participant for this experiment, would you want that person to be included as a participant?": [Yes], [I have no preferences], [No]
- (Q2) "Do you believe that the researchers should be allowed to proceed with this experiment?": [Yes], [Yes, but with caution], [I'm not sure], [No]

Following each question, respondents are provided a space in which they can explain their answer. Finally, respondents—who could be potential participants or people recruited from a micro-task labor markets (e.g., Amazon Mechanical Turk)—answer some basic demographic questions (Schechter and Bravo-Lillo 2014).

Ethical-response surveys have two features that I find particularly attractive. First, they happen before a study has been conducted, and therefore can prevent problems before the research starts (as opposed to approaches that monitor for adverse reactions). Second, ethical-response surveys enable researchers to pose multiple versions of a research project in order to assess the perceived ethical balance of different versions of the same project. One limitation, however, of ethical-response surveys is that it is not clear how to decide between different research designs given the survey results. In cases of extreme uncertainty this kind of information might help guide researchers' decisions; in fact, Schechter and Bravo-Lillo (2014) report abandoning a planned study in response to concerns raised by participants in an ethical-response survey.

While ethical-response surveys can be helpful for assessing reactions to proposed research, they cannot measure the probability or severity of adverse events. One way that medical researchers deal with uncertainty in high-risk settings is *staged trials*, an approach that might be helpful in some social research.

When testing the effectiveness of a new drug, researchers do not immediately jump to a large randomized clinical trial. Rather, they run two types of studies first. Initially, in a Phase I trial, researchers are particularly focused on finding a safe dose, and these studies involve a small number of people. Once a safe dose is discovered, Phase II trials assess the efficacy of the drug, it's ability to work in a best-case situation (Singal, Higgins, and Waljee 2014). Only after Phase I and II studies is a new drug allowed to be assessed in a large randomized controlled trial. While the exact structure of staged trials used in the development of new drugs may not be a good fit for social research, when faced with uncertainty, researchers could run smaller studies explicitly designed to assess safety and efficacy. For example, with Encore, you could imagine the researchers starting with participants in countries with strong rule-of-law.

Together these four approaches—the minimal risk standard, power analysis, ethical-response surveys, and staged trials—can help you proceed in a sensible way, even in the face of uncertainty. Uncertainty need not lead to inaction.

## 1.7 Practical tips

> **In addition to high-minded ethical principles, there are practical issues in research ethics.**

In addition to the ethical principles and frameworks described in this chapter, I'd also like to offer three practical tips based on my personal experience conducting, reviewing, and discussing social research in the digital age.

### 1.7.1 The IRB is a floor, not a ceiling

Many researchers seem to hold contradictory views of the IRB. On the one hand, they consider the IRB to be a bumbling bureaucracy. Yet, at the same time, they also consider the IRB to be the final arbitrator of ethical decisions. That is, they seem to believe that if the IRB approves it, then it must be OK. If we acknowledge the very real limitations of IRBs as they currently exist—and there are many of them (Schrag 2010; Schrag 2011; Hoonaard 2011; Klitzman 2015; King and Sands 2015; Schneider 2015)—then we as researchers must take on additional responsibility for the ethics of our research. The IRB is a floor not a ceiling, and this idea has two main implications.

First, *the IRB is a floor* means that if you are working at an institution that requires IRB review, then you should follow those rules. This may seem obvious, but I've noticed that some people seem to want to avoid the IRB. In fact, if you are working in ethically unsettled areas, the IRB can be a powerful ally. If you follow their rules, they should stand behind you should something go wrong with your research (King and Sands 2015). And, if you don't follow their rules, you could find you are out on your own in a very difficult situation.

Second, *the IRB is not a ceiling* means that just filling out your forms and following the rules in not enough. In many situations you as the researcher are the one who knows the most about how to act ethically. Ultimately, you are the researcher and the ethical responsibility lies with you; it is your name on the paper.

One way to ensure that you treat the IRB as a floor and not a ceiling is to include an ethical appendix in your papers. In fact, you could draft your ethical appendix before your study even begins in order to force yourself to think about how you will explain your work to your peers and the public. If you find yourself uncomfortable while writing your appendix, then your study might not strike the appropriate ethical balance. In addition to helping you diagnose your own work, publishing your ethical appendices will help the research community discuss ethical issues and establish appropriate norms based on examples from real empirical research. Table 3 present empirical research papers that I think have good discussions of research ethics. I don't agree with every claim by the authors in these discussions, but they are all examples of researchers acting with *integrity* in the sense defined by Carter (1996): in each case, the researchers (1) decide what they think is right and what is wrong; (2) they act based on what they have decided, even at personal cost; and (3) they publicly show that they are acting based on their ethical analysis of the situation.

Table 3: Papers with interesting discussions of the ethics of their research.

| Study | Issue addressed |
|---|---|
| Rijt et al. (2014) | field experiments without consent |
| | avoiding contextual harm |
| Paluck and Green (2009) | field experiments in developing country |
| | research on sensitive topic |
| | complex consent issues |
| | remediation of possible harms |
| Burnett and Feamster (2015) | research without consent |
| | balancing risks and benefits when risks are hard to quantify |
| Chaabane et al. (2014) | social implications of research |
| | using leaked data files |
| Jakobsson and Ratkiewicz (2006) | field experiments without consent |
| Soeller et al. (2016) | violated terms of service |

### 1.7.2 Put yourself in everyone else's shoes

Often researchers are so focused on the scientific aims of their work, they see the world only through that lens. This myopia can lead to bad ethical judgment. Therefore, when you are thinking about your study, try to imagine how your participants, other relevant stakeholders, and even a journalist might react to your study. This perspective-taking is different than imaging how *you* would feel in each of these positions. Rather, it is trying to imagine how these *other people* will feel, a processes that is likely to induce empathy (Batson, Early, and Salvarani 1997). Thinking through your work from these different perspectives can help you foresee problems and move your work into better ethical balance.

Further, when imagining your work from the perspective of others, you should expect that they are likely to fixate on vivid worst-case scenarios. For example, in response to Emotional Contagion, some critics focused on the possibility that it might have triggered suicide, a low-probability but extremely vivid worst-case scenario. Once people's emotions are activated and they are focusing on worst-case scenarios, they may completely lose track of the probability of this worst-case event occurring (Sunstein 2002). The fact that people might respond emotionally, however, does not mean that you should dismiss them as uninformed, irrational, or stupid. We should all be humble enough to realize that none of us have the perfect view of ethics.

### 1.7.3 Think of research ethics as continuous, not discrete

Debate about the ethics of social research in the digital age frequently happens in binary terms; for example, Emotional Contagion was either ethical or it was not ethical. This binary thinking polarizes discussion, hinders efforts to develop shared norms, promotes intellectual laziness, and absolves researchers whose research is labeled "ethical" from their responsibility to act more ethically. The most productive conversations that I've seen involving research ethics move beyond this binary thinking to a continuous notion about research ethics.

A major practical problem with binary conceptions of research ethics is that it polarizes discussion. Calling Emotional Contagion or Taste, Ties, and Time unethical lump these studies together with true atrocities in a way that it is not helpful. Moving away from binary thinking and polarizing language is not a call for us to use muddled language to hide unethical behavior. Rather, a continuous notion of ethics will, I think, lead to more careful and precise language. Further, a continuous notion of research ethics clarifies that everyone—even researchers who are doing work that is already considered "ethical"—should strive to create an even better ethical balance in their work.

A final benefit of a move toward continuous thinking is that it encourages intellectual humility, which is appropriate in the face of difficult ethical challenges. The questions of research ethics in the digital age are difficult, and no single person should be so confident in her own ability to diagnose the correct course of action.

## 1.8 Conclusion

Social research in the digital age raises new ethical issues. But, these issues are not insurmountable. If we, as a community, can develop shared ethical norms and standards that are supported both by researchers and the public, then we can harness the capabilities of the digital age in ways that are responsible and beneficial to society. This chapter represents my attempt to move us in that direction, and I think the key will be for researchers to adopt principles-based thinking, while continuing to follow appropriate rules.

In terms of scope, this chapter has focused on the perspective of an individual researcher seeking generalizable knowledge. As such, it leaves out important questions about improvements to the system of ethical oversight of research; questions about regulation of the collection and use of data by companies; and questions about mass surveillance by governments. These other questions are obviously complex and difficult, but it is my hope that some of the ideas from research ethics will be helpful in these other contexts.

## Historical appendix

Any discussion of research ethics needs to acknowledge that, in the past, researchers have done awful things in the name of science. One of the most awful was the Tuskegee Syphilis Study. In 1932, researchers from the US Public Health Service (PHS) enrolled about 400 black men infected with syphilis in a study to monitor the effects of the disease. These men were recruited from the area around Tuskegee, Alabama. From the outset the study was non-therapeutic; it was designed to merely document the history of the disease in black males. The participants were deceived about the nature of the study—they were told that it was a study of "bad blood"—and they were offered false and ineffective treatment, even though syphilis is a deadly disease. As the study progressed, safe and effective treatments for syphilis were developed, but the researchers actively intervened to prevent the participants from getting treatment elsewhere. For example, during World War II the research team secured draft deferments for all men in the study in order to prevent the treatment the men would have received had they entered the Armed Forces. Researchers continued to deceive participants and deny them care for 40 years. The study was a 40-year deathwatch.

The Tuskegee Syphilis Study took place against a backdrop of racism and extreme inequality that was common in the southern part of the US at the time. But, over its 40-year history, the study involved dozens of researchers, both black and white. And, in addition to researchers directly involved, many more must have read one of the 15 reports of the study published in the medical literature (Heller 1972). In the mid-1960s—about 30 years after the study began—a PHS employee named Robert Buxtun began pushing within the PHS to end the study, which he considered morally outrageous. In response to Buxtun, in 1969 the PHS convened a panel to do a complete ethical review of the study. Shockingly, the ethical review panel decided that researchers should continue to withhold treatment from the infected men. During the deliberations, one member of the panel even remarked: "You will never have another study like this; take advantage of it" (Brandt 1978). The all white panel, which was mostly made up of doctors, did decide that some form of informed consent should be acquired. But, the panel judged the men themselves incapable of providing informed consent because of their age and low level of education. The panel recommended, therefore, that the researchers receive "surrogate informed consent" from local medical officials. So, even after an a full ethical review, the withholding of care continued. Eventually, Robert Buxtun took the story to a journalist, and in 1972 Jean Heller wrote a series of newspaper articles that exposed the study to the world. It was only after widespread public outrage that the study was finally ended and care was offered to the men who had survived.

Table 4: Partial time line of the Tuskegee Syphilis Study, adapted from Jones (2011).

| Date | Event |
|---|---|
| 1932 | approximately 400 men with syphilis are enrolled in the study; they are not informed of the nature of the research |
| 1937-38 | PHS sends mobile treatment units to area, but treatment is withheld for men in study |
| 1942-43 | PHS intervenes to prevent men from being drafted for WWII in order to prevent them from receiving treatment |
| 1950s | Penicillin becomes a widely available and effective treatment for syphilis; men are still not treated (Brandt 1978) |
| 1969 | PHS convenes an ethical review of the study; panel recommends that the study continue |
| 1972 | Peter Buxtun, a former PHS employee, tells a reporter about the study; and press breaks the story |
| 1972 | US Senate holds hearings on human experimentation, including Tuskegee Study |
| 1973 | Government officially ends the study and authorizes treatment for survivors |
| 1997 | US President Bill Clinton publicly and officially apologizes for the Tuskegee Study |

Victims of this study include not just the 399 men, but also their families: at least 22 wives, 17 children, and 2 grandchildren with syphilis may have contracted the disease as a result of the withholding of treatment (Yoon 1997). Further, the harm caused by the study continued long after it ended. The study—justifiably—decreased the trust that African Americans had in the medical community, an erosion in trust that may have led African-Americans to avoid medical care to the determent of their health (Alsan and Wanamaker 2016). Further, the lack of trust hindered efforts to treat HIV/AIDS in the 1980s and 90s (Jones 1993, Ch. 14).

Although it is hard to imagine research so horrific happening today, I think there are three important lessons from the Tuskegee Syphilis Study for people conducting social research in the digital age. First, it reminds us that there are some studies that simply should not happen. Second, it shows us that research can harm not just participants, but also their families and entire communities long after the research has been completed. Finally, it shows that researchers can make terrible ethical decisions. In fact, I think it should induce some fear in researchers today that so many people involved in this study made such awful decisions over such a long period of time. And, unfortunately, Tuskegee is by no means unique; there were several other examples of problematic social and medical research during this era (Katz, Capron, and Glass 1972; Emanuel et al. 2008).

In 1974, in response to the Tuskegee Syphilis Study and these other ethical failures by researchers, the US Congress created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research and tasked the committee to develop ethical guidelines for research involving human subjects. After four years of meeting at the Belmont Conference Center, the group produced the *Belmont Report*, a slender but powerful document that has had a tremendous impact on both abstract debates in bioethics and the everyday practice of research.

The Belmont Report has three sections. In the first section—Boundaries Between Practice and Research—the Belmont Report sets out its purview. In particular, it argues for a distinction between *research*, which seeks generalizable knowledge, and *practice*, which includes everyday treatment and activities. Further, it argues that the ethical principles of the Belmont Report apply only to research. It has been argued that this distinction between research and practice is one way that the Belmont Report is misfit to social research in the digital age (Metcalf and Crawford 2016; boyd 2016).

The second and third parts of the Belmont Report lay out three ethical principles—Respect for Persons; Beneficence; and Justice—and describe how these principles can be applied in research practice. These are the principles that I described in more detail in the chapter.

The Belmont Report sets broad goals, but it is not a document that can be easily used to oversee day-to-day activities. Therefore, the US Government created a set of regulations that are colloquially called the *Common Rule* (their official name is Title 45 Code of Federal Regulations, Part 46, Subparts A - D) (Porter and Koski 2008). These regulations describe the process for reviewing, approving, and overseeing research, and they are the regulations that Institutional Review Boards (IRBs) are tasked with enforcing. To understand the difference between the Belmont Report and the Common Rule, consider how each discusses informed consent: the Belmont Report describes the philosophical reasons for informed consent and broad characteristics that would represent true informed consent while the Common Rule lists the eight required and six optional elements of an informed consent document. By law, the Common Rule governs almost all research that receives funding from the US Government. Further, many institutions that receive funding from the US Government typically apply the Common Rule to all research happening at that institution, regardless of the funding source. But, the Common Rule does not automatically apply at companies that do not receive research funding from the US Government.

I think that almost all researchers respect the broad goals of ethical research as expressed in the Belmont Report, but there is widespread annoyance with the Common Rule and the process of working with IRBs (Schrag 2010; Schrag 2011; Hoonaard 2011; Klitzman 2015; King and Sands 2015; Schneider 2015). To be clear, those critical of IRBs are not against ethics. Rather, they believe that the current system does not strike an appropriate balance or could better achieve its goals through other methods. This chapter, however, will take these IRBs as given. If you are required to follow the rules of an IRB, then you should follow them. However, I would encourage you to *also* take a principles-based approach when considering the ethics of your research.

This background very briefly summarizes how we arrived at the rules-based system of IRB review in the United States. When considering the Belmont Report and the Common Rule today, we should remember that they were created in a different era and were—quite sensibly—responding to the problems of that era, in particular breaches in medical ethics during and after the Second World War (Beauchamp 2011).

In addition to ethical efforts by medical and behavioral scientists to create ethical codes, there were also smaller and less well known efforts by computer scientists. In fact, the first researchers to run into the ethical challenges created by digital age research were not social scientists; they were computer scientists, specifically researchers in computer security. During the 1990s and 2000s computer security researchers conducted a number of ethically questionable studies that involved things like taking over botnets and hacking into thousands of computers with weak passwords (Bailey, Dittrich, and Kenneally 2013; Dittrich, Carpenter, and Karir 2015). In response to these studies, the US Government—specifically the Department of Homeland Security—created a blue-ribbon commission to write a guiding ethical framework for research involving information and communication technologies (ICT). The results of this effort was the *Menlo Report* (Dittrich, Kenneally, and others 2011). Although the concerns of computer security researchers are not exactly the same as social researchers, the Menlo Report provides three important lessons for social researchers.

First, the Menlo Report reaffirms the three Belmont principles—Respect for Persons, Beneficence, and Justice—and adds a fourth principle: *Respect for Law and Public Interest*. I described this fourth principle and how it should be applied to social research in the main chapter (Section 1.4.4).

Second, the Menlo Report calls on researchers to move beyond a narrow definition of "research involving human subjects" from the Belmont Report to a more general notion of "research with human-harming potential." The limitations of the scope of the Belmont Report are well illustrated by Encore. The IRBs at Princeton and Georgia Tech ruled that Encore was not "research involving human subjects," and therefore not subject to review under the Common Rule. However, Encore clearly has human-harming potential; at its most extreme, Encore could potentially result in innocent people being jailed by repressive governments. A principles-based approaches mean that researchers should not hide behind a narrow, legal definition of "research involving human subjects," even if IRBs allow it. Rather, they should adopt a more general notion of "research with human-harming potential" and they should subject all of their own research with human-harming potential to ethical consideration.

Third, the Menlo Report calls on researchers to expand the stakeholders that are considered when applying the Belmont principles. As research has moved from a separate sphere of life to something that is more embedded in day-to-day activities, ethical considerations must be expanded beyond just specific research participants to include non-participants and the environment where the research takes place. In other words, the Menlo Report calls for researchers to broaden their ethical field of view beyond just their participants.

This historical appendix provides a very brief review of research ethics in the social and medical science, as well as computer science. For a book length treatment of research ethics in medical science, see Emanuel et al. (2008) or Beauchamp and Childress (2012).

## Further commentary

**This section is designed to be used as a reference, rather than to be read as a narrative.**

- **Introduction (Section 1.1)**

Research ethics has traditionally also included topics such as scientific fraud and allocation of credit. These topics are discussed in greater detail in Engineering (2009).

This chapter is strongly shaped by the situation in the United States. For more on the ethical review procedures in other countries, see Chapters 6, 7, 8, and 9 of Desposato (2016b). For an argument that the biomedical ethical principles that have influenced this chapter are excessively American, see Holm (1995). For more historical review of Institutional Review Boards in the US, see Stark (2012).

The Belmont Report and subsequent regulations in the US have made a distinction between research and practice. This distinction has been criticized subsequently (Beauchamp and Saghai 2012; boyd 2016; Metcalf and Crawford 2016; Meyer 2015). I do not make this distinction in this chapter because I think the ethical principles and frameworks apply to both settings. For more on research oversight at Facebook, see Jackman and Kanerva (2016). For a proposal for research oversight at companies and NGOs, see Polonetsky, Tene, and Jerome (2015) and Tene and Polonetsky (2016).

For more on the case of the Ebola outbreak in 2014, see McDonald (2016), and for more about the privacy risks of mobile phone data, see Mayer, Mutchler, and Mitchell (2016). For an example of crisis-related research using mobile phone data, see Bengtsson et al. (2011) and Lu, Bengtsson, and Holme (2012).

- **Three examples (Section 1.2)**

Many people have written about Emotional Contagion. The journal *Research Ethics* devoted their entire issue in January 2016 discussing the experiment; see Hunter and Evans (2016) for an overview. The *Proceedings of the National Academics of Science* published two pieces about the experiment: Kahn, Vayena, and Mastroianni (2014) and Fiske and Hauser (2014). Other pieces about the experiment include: Puschmann and Bozdag (2014); Meyer (2014); Grimmelmann (2015); Meyer (2015); Selinger and Hartzog (2015); Kleinsman and Buckley (2015); Shaw (2015); Flick (2015).

For more on Encore, see Jones and Feamster (2015).

- **Digital is different (Section 1.3)**

In terms of mass surveillance, broad overviews are provided in Mayer-Schönberger (2009) and Marx (2016). For a concrete example of the changing costs of surveillance, Bankston and Soltani (2013) estimates that tracking a criminal suspect using cell phones is about 50 times cheaper than using physical surveillance. Bell and Gemmell (2009) provides a more optimistic perspective on self-surveillance. In addition to being able to track observable behavior that is public or partially public (e.g., Taste, Ties, and Time), researchers can increasingly infer things that many participants consider to be private. For example, Michal Kosinski and colleagues showed that they could infer sensitive information about people, such as sexual orientation and use of addictive substances from seemingly ordinary digital trace data (Facebook Likes) (Kosinski, Stillwell, and Graepel 2013). This might sound magical, but the approach Kosinski and colleagues used—which combines digital traces, surveys, and supervised learning—is actually something that I've already told you about. Recall that in Chapter 3 (Asking questions) I told you how Josh Blumenstock and colleagues (2015) combined survey data with mobile phone data to estimate poverty in Rwanda. This exact same approach, which can be used to efficiently measure poverty in a developing country, can also be used for potentially privacy violating inferences.

Inconsistent laws and norms can lead to research that does not respect the wishes of participants, and it can lead to "regulatory shopping" by researchers (Grimmelmann 2015; Nickerson and Hyde 2016). In particular, some researchers who wish to avoid IRB oversight have partners who are not covered by IRBs (e.g., people at companies or NGOs) collect and de-identify data. Then, the researchers can analyze this de-identified data without IRB oversight, at least according to some interpretations of current rules. This kind of IRB evasion appears to be inconsistent with a principles-based approach.

For more on the inconsistent and heterogeneous ideas that people have about health data, see Fiore-Gartland and Neff (2015). For more on the problem that heterogeneity creates for research ethics decisions see Meyer (2013).

One difference between analog age and digital age research is that in digital age research *interaction with participants is more distant*. These interactions often occur through an intermediary such as a company, and there is typically a large physical—and social—distance between researchers and participants. This distant interaction makes some things that are easy in analog age research difficult in digital age research, such as screening out participants who require extra protection, detecting adverse events, and remediating harm if it occurs. For example, let's contrast Emotional Contagion with a hypothetical lab experiment on the same topic. In the lab experiment, researchers could screen out anyone who arrives at the lab showing obvious signs of emotional distress. Further, if the lab experiment created an adverse event, researchers would see it, provide services to remediate the harm, and then make adjustments to the experimental protocol to prevent future harms. The distant nature of interaction in the actual Emotional Contagion experiment makes each of these simple and sensible steps extremely difficult. Also, I suspect that the distance between researchers and participants makes researchers less sensitive to the concerns of their participants.

Other sources of inconsistent norms and laws. Some of this inconsistency comes from the fact that this research is happening all over the world. For example, Encore involved people from all over the world, and therefore it might be subject to the data protection and privacy laws of many different countries. What if the norms governing third-party web requests (what Encore was doing) are different in Germany, the United States, Kenya, and China? What if the norms are not even consistent within a single country? A second source of inconsistency comes from collaborations between researchers at universities and companies; for example, Emotional Contagion was a collaboration between a data scientist at Facebook and a professor and graduate student at Cornell. At Facebook running large experiments is routine and, at that time, did not require any third-party ethical review. At Cornell the norms and rules are quite different; virtually all experiments must be reviewed by the Cornell IRB. So, which set of rules should govern Emotional Contagion—Facebook's or Cornell's?

For more on efforts to revise the Common Rule, see Evans (2013),Council (2014), Metcalf (2016), and Hudson and Collins (2015).

- **Four principles (Section 1.4)**

The classic principles-based approach to biomedical ethics is Beauchamp and Childress (2012). They propose that four main principles should guide biomedical ethics: Respect for Autonomy, Nonmaleficence, Beneficence, and Justice. The principle of nonmaleficence urges one to abstain from causing harm to other people. This concept is deeply connected to Hippocratic idea of "Do no harm." In research ethics, this principle is often combined with the principle of Beneficence, but see Beauchamp and Childress (2012) (Chapter 5) for more on the distinction between the two. For a criticism that these principles are overly American, see Holm (1995). For more on balancing when the principles conflict, see Gillon (2015).

The four principles in this chapter have also been proposed to guide ethical oversight for research happening at companies and NGOs (Polonetsky, Tene, and Jerome 2015) through bodies called "Consumer Subject Review Boards" (CSRBs) (Calo 2013).

- **Respect for Persons (Section 1.4.1)**

In addition to respecting autonomy, the Belmont Report also acknowledges that not every human is capable of true self-determination. For example, children, people suffering from illness, or people living in situations of severely restricted liberty may not be able to act as fully autonomous individuals, and these people are, therefore, subject to extra protection.

Applying the principle of Respect for Persons in the digital age can be challenging. For example, in digital age research, it can be difficult to provide extra protections for people with diminished capability of self-determination because researchers often know very little about their participants. Further, informed consent in digital age social research is a huge challenge. In some cases, truly informed consent can suffer from the *transparency paradox* (Nissenbaum 2011), where *information* and *comprehension* are in conflict. Roughly, if researchers provide full information about the nature of the data collection, data analysis, and data security practices, it will be difficult for many participants to comprehend. But, if researchers provide comprehensible information, it may lack important technical information. In medical research in the analog age—the dominate setting considered by the Belmont Report—one could imagine a doctor talking individually with each participant to help resolve the transparency paradox. In online studies involving thousands or millions of people, such a face-to-face approach is impossible. A second problem with consent in the digital age is that in some studies, such as analysis of massive data repositories, it would be impractical to obtain informed consent from all participants. I discuss these and other questions about informed consent in more detail in Section 1.6.1. Despite these difficulties, however, we should remember that informed consent is neither necessary nor sufficient for Respect for Persons.

For more on medical research before informed consent, see Miller (2014). For a book-length treatment of informed consent, see Manson and O'Neill (2007). See also the suggested readings about informed consent below.

- **Beneficence (Section 1.4.2)**

Harms to context is the harm that research can cause not to specific people but to social settings. This concept is a bit abstract, but I'll illustrate it with two examples: one analog and one digital.

A classic example of harms to context comes from the Wichita Jury Study [Vaughan (1967); Katz, Capron, and Glass (1972); Ch 2.]—also sometimes called the Chicago Jury Project (Cornwell 2010). In this study researchers from the University of Chicago, as part of a larger study about social aspects of the legal system, secretly recorded six jury deliberations in Wichita, Kansas. The judges and lawyers in the cases had approved the recordings, and there was strict oversight of the process. However, the jurors were unaware that recordings were occurring. Once the study was discovered, there was public outrage. The Justice Department began an investigation of the study, and the researchers were called to testify in front of Congress. Ultimately, Congress passed a new law that makes it illegal to secretly record jury deliberation.

The concern of critics of the Wichita Jury Study was not harm to participants; rather, it was harms to the context of jury deliberation. That is, people believed that if jury members did not believe that they were having discussions in a safe and protected space, it would be harder for jury deliberations to proceed in the future. In addition to jury deliberation, there are other specific social contexts that society provides with extra protection such as attorney-client relationships and psychological care (MacCarthy 2015).

The risk of harms to context and the disruption of social systems also comes up in some field experiments in Political Science (Desposato 2016b). For an example of a more context-sensitive cost-benefit calculation for a field experiment in Political Science, see Zimmerman (2016).

- **Justice (Section 1.4.3)**

Compensation for participants has been discussed in a number of settings related to digital age research. Lanier (2014) proposed paying participants for digital traces they generate. Bederson and Quinn (2011) discusses payments in online labor markets. Finally, Desposato (2016a) proposes paying participants in field experiments. He points out that even if participants cannot be paid directly, a donation could be made to a group working on their behalf. For example, in Encore the researchers could have made a donation to a group working to support access to the Internet.

- **Respect for Law and Public Interest (Section 1.4.4)**

Terms-of-service agreements should have less weight than contracts negotiated between equal parties and laws created by legitimate governments. Situations where researchers have violated terms-of-service agreements in the past generally involve using automated queries to audit the behavior of companies (much like field experiments to measure discrimination). For additional discussion see Vaccaro et al. (2015), Bruckman (2016a), Bruckman (2016b). For an example of empirical research that discusses terms of service, see Soeller et al. (2016). For more on the possible legal problems researchers face if they violate terms of service see Sandvig and Karahalios (2016).

- **Two ethical frameworks (Section 1.5)**

Obviously, enormous amounts have been written about consequentialism and deontology. For an example of how these ethical frameworks, and others, can be used to reason about digital age research, see Zevenbergen et al. (2015). For an example of how these ethical frameworks can be applied to field experiments in develop economics, see Baele (2013).

- **Informed consent (Section 1.6.1)**

For more on audit studies of discrimination, see Pager (2007) and Riach and Rich (2004). Not only do these studies not have informed consent, they also involve deception without debriefing.

Both Desposato (2016a) and Humphreys (2015) offer advice about field experiments without consent.

Sommers and Miller (2013) reviews many arguments in favor of not debriefing participants after deception, and argues that researchers should forgo "debriefing under a very narrow set of circumstances, namely, in field research in which debriefing poses considerable practical barriers but researchers would have no qualms about debriefing if they could. Researchers should not be permitted to forgo debriefing in order to preserve a naive participant pool, shield themselves from participant anger, or protect participants from harm." Others argue that if debriefing causes more harm than good it should be avoided. Debriefing is a case where some researchers prioritize Respect for Persons over Beneficence, and some researchers do the opposite. One possible solution would be to find ways to make debriefing a learning experience for the participants. That is, rather than thinking of debriefing as something that can cause harm, perhaps debriefing can also be something that benefits participants. For an example of this kind of education debriefing, see Jagatic et al. (2007) on debriefing students after a social phishing experiment. Psychologists have developed techniques for debriefing (Holmes 1976a; Holmes 1976b; Mills 1976; Baumrind 1985; Oczak and Niedźwieńska 2007) and some of these may be usefully applied to digital age research. Humphreys (2015) offers interesting thoughts about *deferred consent*, which is closely related to the debriefing strategy that I described.

The idea of asking a sample of participants for their consent is related to what Humphreys (2015) calls *inferred consent.*

A further idea that has been proposed related to informed consent is to build a panel of people who agree to be in online experiments (Crawford 2014). Some have argued that this panel would be a non-random sample of people. But, Chapter 3 (Asking questions) shows that these problems are potentially addressable using post-stratification and sample matching. Also, consent to be on the panel could cover a variety of experiments. In other words, participants might not need to consent to each experiment individually, a concept called *broad consent* (Sheehan 2011).

- **Understanding and managing informational risk (Section 1.6.2)**

Far from unique, the Netflix Prize illustrates an important technical property of datasets that contain detailed information about people, and thus offers important lessons about the possibility of "anonymization" of modern social datasets. Files with many pieces of information about each person are likely to be *sparse*, in the sense defined formally in Narayanan and Shmatikov (2008). That is, for each record there are no records that are the same, and in fact there are no records that are very similar: each person is far away from their nearest neighbor in the dataset. One can imagine that the Netflix data might be sparse because with about 20,000 movies on a 5 star scale, there are about $6^{20,000}$ possible values that each person could have (6 because in addition to one to 5 stars, someone might have not rated the movie at all). This number is so large, it is hard to even comprehend.

Sparsity has two main implications. First, it means that attempting to "anonymize" the dataset based on random perturbation will likely fail. That is, even if Netflix were to randomly adjust some of the ratings (which they did), this would not be sufficient because the perturbed record is still the closest possible record to the information that the attacker has. Second, the sparsity means that de-anonymization is possible even if the attacker has imperfect or impartial knowledge. For example, in the Netflix data, let's imagine the attacker knows your ratings for two movies and the dates you made those ratings +/- 3 days; just that information alone is sufficient to uniquely identify 68% of people in the Netflix data. If the attackers knows 8 movies that you have rated +/- 14 days, then even if two of these known ratings are completely wrong, 99% of records can be uniquely identified in the dataset. In other words, sparsity is a fundamental problem for efforts to "anonymize" data, which is unfortunate because most modern social dataset are sparse.

Telephone metadata also might appear to be "anonymous" and not sensitive, but that is not the case. Telephone metadata is identifiable and sensitive (Mayer, Mutchler, and Mitchell 2016; Landau 2016).

In Figure 6, I sketched out a trade-off between risk to participants and benefits to research from data release. For a comparison between restricted access approaches (e.g., a walled garden) and restricted data approaches (e.g., some form of anonymization) see Reiter and Kinney (2011). For a proposed categorization system of risk levels of data, see Sweeney, Crosas, and Bar-Sinai (2015). Finally, for a more a general discussion of data sharing, see Yakowitz (2011).

For more detailed analysis of this trade-off between the risk and utility of data, see Brickell and Shmatikov (2008), Ohm (2010), Wu (2013), Reiter (2012), and Goroff (2015). To see this trade-off applied to real data from massively open online courses (MOOCs), see Daries et al. (2014) and Angiuli, Blitzstein, and Waldo (2015).

Differential privacy also offers an alternative approach that can combine both high benefit to society and low risk to participants, see Dwork and Roth (2014) and Narayanan, Huey, and Felten (2016).

For more on the concept of personally identifying information (PII), which is central to many of the rules about research ethics, see Narayanan and Shmatikov (2010) and Schwartz and Solove (2011). For more on all data being potentially sensitive, see Ohm (2015).

In this section, I've portrayed the linkage of different datasets as something that can lead to informational risk. However, it can also create new opportunities for research, as argued in Currie (2013).

For more on the five safes, see Desai, Ritchie, and Welpton (2016). For an example of how outputs can be identifying, see Brownstein, Cassa, and Mandl (2006), which shows how maps of disease prevalence can be identifying. Dwork et al. (2017) also considers attacks against aggregate data, such as statistics about how many individuals have a certain disease.

- **Privacy (Section 1.6.3)**

Warren and Brandeis (1890) is a landmark legal article about privacy, and the article is most associated with the idea that privacy is a right to be left alone. More recently book length treatments of privacy that I would recommend include Solove (2010) and Nissenbaum (2010).

For a review of empirical research on how people think about privacy, see Acquisti, Brandimarte, and Loewenstein (2015). The journal *Science* published a special issue titled "The End of Privacy", which addresses the issues of privacy and information risk from a variety of different perspectives; for a summary see Enserink and Chin (2015). Calo (2011) offers a framework for thinking about the harms that come from privacy violations. An early example of concerns about privacy in the very beginnings of the digital age is Packard (1964).

- **Making decisions under uncertainty (Section 1.6.4)**

One challenge when trying to apply the minimal risk standard is that it is not clear whose daily life is to be used for benchmarking (Council 2014). For example, homeless people have higher levels of discomfort in their daily lives. But, that does not imply that it is ethically permissible to expose homeless people to higher risk research. For this reason, there seems to be a growing consensus that minimal risk should be benchmarked against a *general population* standard, not a *specific population* standard. While I generally agree with the idea of a general population standard, I think that for large online platforms such as Facebook, a specific population standard is reasonable. That is, when considering Emotional Contagion, I think that it is reasonable to benchmark against everyday risk on Facebook. A specific population standard in this case is much easier to evaluate and is unlikely to conflict with the principle of Justice, which seeks to prevent the burdens of research failing unfairly on disadvantaged groups (e.g., prisoners and orphans).

- **Practical tips (Section 1.7)**

Other scholars have also called for more papers to include ethical appendices (Schultze and Mason 2012; Kosinski et al. 2015). King and Sands (2015) also offers practical tips.

## Activities

Key:

- degree of difficulty: easy ◔, medium ◑, hard ◕, very hard ●
- requires math (🖩)
- requires coding (🗎)
- data collection (⛁)
- my favorites (♥)

1. [◔] In arguing against the Emotional Contagion experiment, Kleinsman and Buckley (2015) wrote:

   > "Even if it is true that the risks for the Facebook experiment were low and even if, in hindsight, the results are judged to be useful, there is an important principle at stake here that must be upheld. In the same way that stealing is stealing no matter what amounts are involved, so we all have a right not to be experimented on without our knowledge and consent, whatever the nature of the research."

   a) Which of the two ethical frameworks discussed in this chapter—consequentialism or deontology—is this argument most clearly associated with?

   b) Now, imagine that you wanted to argue against this position. How would you argue the case to a reporter for *The New York Times*?
   c) How, if at all, would your argument be different if you were discussing this with a colleague?

2. [◠] Maddock, Mason, and Starbird (2015) considers the question of whether researchers should use tweets that have been deleted. Read their paper to learn about the background.

   a) Analyze this decision from deontological perspective.
   b) Analyze the exact same decision from a consequentialist perspective.
   c) Which do you find more convincing in this case?

3. [◠] In an article on the ethics of field experiments, Humphreys (2015), proposed the following hypothetical experiment to highlight the ethical challenges of interventions that are done without consent of all impacted parties and that harms some and help others.

   > "Say a researcher is contacted by a set of community organizations that want to figure out whether placing street lights in slums will reduce violent crime. In this research the subjects are the criminals: seeking informed consent of the criminals would likely compromise the research and it would likely not be forthcoming anyhow (violation of respect for persons); the criminals will likely bear the costs of the research without benefiting (violation of justice); and there will be disagreement regarding the benefits of the research – if it is effective, the criminals in particular will not value it (producing a difficulty for assessing benevolence). . . . The special issues here are not just around the subjects however. Here there are also risks that obtain to non-subjects, if for example criminals retaliate against the organizations putting the lamps in place. The organization may be very aware of these risks but be willing to bear them because they erroneously put faith in the ill-founded expectations of researchers from wealthy universities who are themselves motivated in part to publish."

   a) Write an email to the community organization offering your ethical assessment of the experiment as designed? Would you help them do the experiment as proposed? What factors might impact your decision?
   b) Are there some changes that might improve your assessment of the ethics of this experimental design.

4. [◠] In the 1970's 60 men participated in field experiment that took place in the men's bathroom at a university in the midwestern part of the US (the researchers don't name the university) (Middlemist, Knowles, and Matter 1976). The researchers were interested in how people respond to violations of their personal space, which Sommer (1969) defined as the "area with invisible boundaries surrounding a person's body into which intruders may not come." More specifically, the researchers chose to study how a man's urination was impacted by the presence of others nearby. After conducting a purely observation study, the researchers conducted a field experiment. Participants were forced to use the left-most urinal in a three urinal bathroom (the researchers do not explain exactly how this happened). Next, participants were assigned to one of three levels of interpersonal distance. For some men a confederate used a urinal right next to them, for some men a confederate used a urinal one space away from them, and for some men no confederate entered the bathroom. The researchers measured their outcome variables—delay time and persistence—by stationing a research assistant inside the toilet stall adjacent to the participant's urinal. Here's how the researchers described the measurement procedure:

   > "An observer was stationed in the toilet stall immediately adjacent to the subjects' urinal. During pilot tests of these procedures it became clear that auditory cues

> could not be used to signal the initiation and cessation of [urination]. . . . Instead, visual cues were used. The observer used a periscopic prism imbedded in a stack of books lying on the floor of the toilet stall. An 11-inch (28-cm) space between the floor and the wall of the toilet stall provided a view, through the periscope, of the user's lower torso and made possible direct visual sightings of the stream of urine. The observer, however, was unable to see a subject's face. The observer started two stop watches when a subject stepped up to the urinal, stopped one when urination began, and stopped the other when urination was terminated."

The researchers found that decreased physical distance leads to increased delay of onset and decreased persistence (Figure 7).

a) Do you think the participants were harmed by this experiment?
b) Do you think that the researchers should have conducted this experiment?
c) What changes, if any, would you recommend to improve the ethical balance?

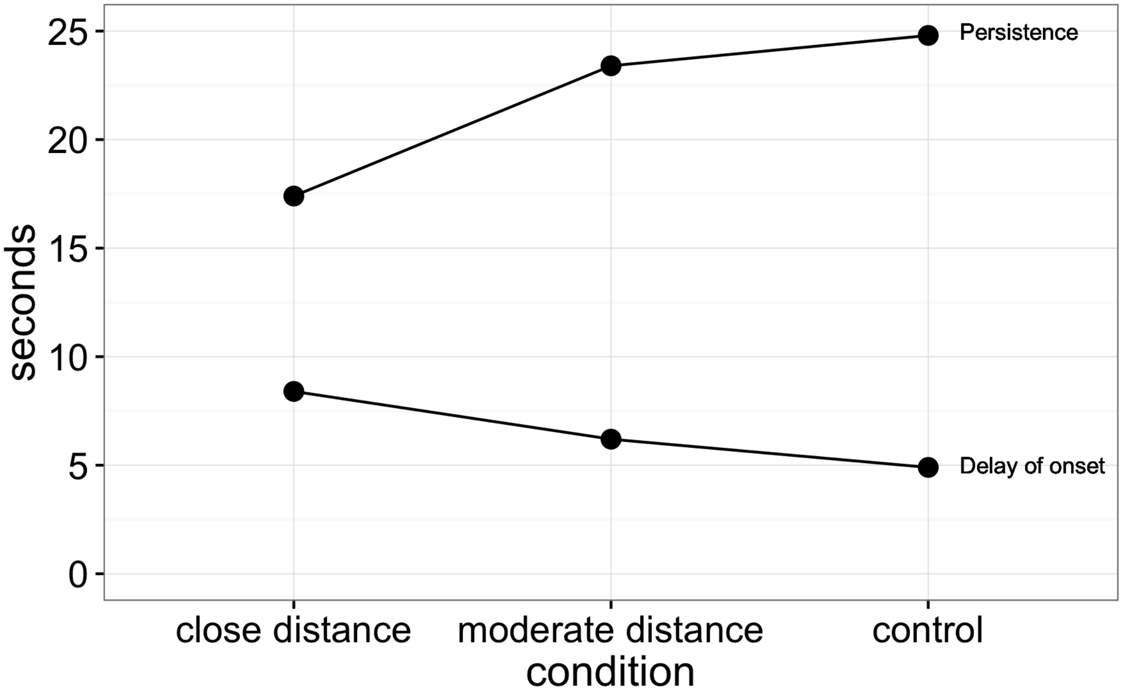


Figure 7: Results from Middlemist, Knowles, and Matter (1976). Men who entered the bathroom were assigned to one of three conditions: close distance (a confederate was placed in the immediately adjacent urinal), moderate distance (a confederate was placed one urinal removed), or no confederate used a urinal. An observer stationed in a toilet stall used a custom-built periscope to observe and time the delay and persistence of urination. Standard errors around estimates are not available.

5. [⏲] In August 2006, about 10 days prior to the a primary election, 20,000 people living in Michigan received a mailing that showed their voting behavior and the voting behavior of their neighbors (Figure 8). (As discussed in the chapter, in the US, state governments keeps records of who votes in each election and this information is available to the public.) This particular treatment produced the largest effect ever seen up to that point for a single piece mailing: it increased the turnout rate by 8.1 percentage points (Gerber, Green, and Larimer 2008). To

put this in context, one piece mailings typically produce increases of about one percentage point (Gerber, Green, and Larimer 2008). The effect was so large that a political operative named Hal Malchow offered Donald Green $100,000 not to publish the result of the experiment (presumably so that Malchow could make use of this information himself) (Issenberg 2012, p 304). But, Alan Gerber, Donald Green, and Christopher Larimer did publish the paper in 2008 in the *American Political Science Review.*

When you carefully inspect the mailer in Figure 8 you may notice that the researchers' names do not appear on it. Rather, the return address is to Practical Political Consulting. In the acknowledgment to the paper the authors explain: "Special thanks go to Mark Grebner of Practical Political Consulting, who designed and administered the mail program studied here."

a) Please assess the use of this treatment in terms of the four ethical principles described in this chapter.
b) What changes, if any, would you recommend to this experiment?
c) Write an ethical appendix that could appear with this paper when it was published.

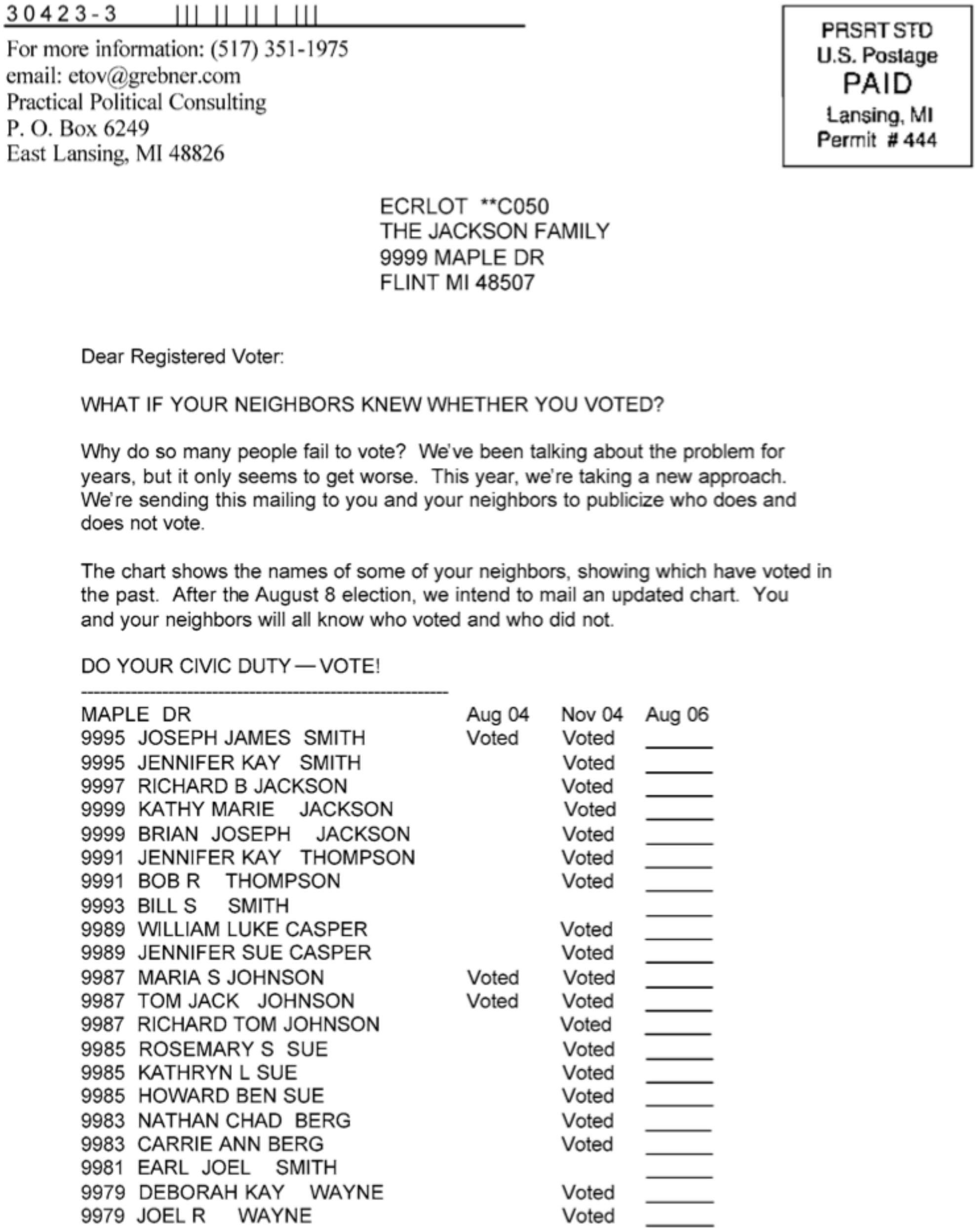
6. [◠] Building on the previous question, once these 20,000 mailers were sent (Figure 8), as well as 60,000 other potentially less sensitive mailers, there was a backlash from participants. In fact, Issenberg (2012) (p 198) reports that "Grebner [the director of Practical Political Consulting] was never able to calculate how many people took the trouble to complain by phone, because his office answering machine filled so quickly that new callers were unable to leave a message." In fact, Grebner noted that the backlash could have been even larger if they had scaled up the treatment. He said to Alan Gerber, one of the researchers, "Alan if we had spent five hundred thousand dollars and covered the whole state you and I would be living with Salman Rushdie." (Issenberg 2012, p 200)

   a) Does this information change your answers to the previous question?
   b) What strategies for dealing with making decisions in the face of uncertainty would you recommend for future studies that are similar?

7. [◠] In practice, most ethical debate occurs about studies where researchers do not have true informed consent from participants (e.g., the three case studies in this chapter). However, ethical debate can also occur for studies that have true informed consent. Design a hypothetical study where you would have true informed consent from participants, but which you still think would be unethical. (Hint: If you are struggling, you can try reading Emanuel, Wendler, and Grady (2000).)

8. [◠] Researchers often struggle to describe their ethical thinking to each other and to the general public. After it was discovered that Taste, Ties, and Time was re-identified, Jason Kauffman, the leader of the research team, made a few public comments about the ethics of the project. Read Zimmer (2010) and then rewrite Kauffman's comments using the principles and ethical frameworks that are described in this chapter.

9. [◠] Banksy is one of the most famous contemporary artist in the United Kingdom, and he is know for politically-oriented street graffiti (Figure 9). His precise identity, however, is a mystery. Banksy has a personal website so he could make his identity public if he wanted, but he has chosen not to. In 2008 the *Daily Mail*, a newspaper, published an article claiming to identify Banksy's real name. Then in 2016, Michelle Hauge, Mark Stevenson, D. Kim Rossmo and Steven C. Le Comber (2016) attempted to verify this claim using Dirichlet process mixture model of geographic profiling. More specifically, they collected the geographic locations of Banksy's public graffiti in Bristol and London. Next, by searching through old newspaper articles and public voting records, Hauge and colleagues found past addresses of the named individual, his wife, and his football (i.e., soccer) team. The author's summarize the finding of their paper as follows:



30423-3

For more information: (517) 351-1975
email: etov@grebner.com
Practical Political Consulting
P. O. Box 6249
East Lansing, MI 48826

PRSRT STD
U.S. Postage
PAID
Lansing, MI
Permit # 444

ECRLOT **C050
THE JACKSON FAMILY
9999 MAPLE DR
FLINT MI 48507

Dear Registered Voter:

WHAT IF YOUR NEIGHBORS KNEW WHETHER YOU VOTED?

Why do so many people fail to vote? We've been talking about the problem for years, but it only seems to get worse. This year, we're taking a new approach. We're sending this mailing to you and your neighbors to publicize who does and does not vote.

The chart shows the names of some of your neighbors, showing which have voted in the past. After the August 8 election, we intend to mail an updated chart. You and your neighbors will all know who voted and who did not.

DO YOUR CIVIC DUTY — VOTE!

| MAPLE DR | Aug 04 | Nov 04 | Aug 06 |
|---|---|---|---|
| 9995 JOSEPH JAMES SMITH | Voted | Voted | ______ |
| 9995 JENNIFER KAY SMITH | | Voted | ______ |
| 9997 RICHARD B JACKSON | | Voted | ______ |
| 9999 KATHY MARIE JACKSON | | Voted | ______ |
| 9999 BRIAN JOSEPH JACKSON | | Voted | ______ |
| 9991 JENNIFER KAY THOMPSON | | Voted | ______ |
| 9991 BOB R THOMPSON | | Voted | ______ |
| 9993 BILL S SMITH | | | ______ |
| 9989 WILLIAM LUKE CASPER | | Voted | ______ |
| 9989 JENNIFER SUE CASPER | | Voted | ______ |
| 9987 MARIA S JOHNSON | Voted | Voted | ______ |
| 9987 TOM JACK JOHNSON | Voted | Voted | ______ |
| 9987 RICHARD TOM JOHNSON | | Voted | ______ |
| 9985 ROSEMARY S SUE | | Voted | ______ |
| 9985 KATHRYN L SUE | | Voted | ______ |
| 9985 HOWARD BEN SUE | | Voted | ______ |
| 9983 NATHAN CHAD BERG | | Voted | ______ |
| 9983 CARRIE ANN BERG | | Voted | ______ |
| 9981 EARL JOEL SMITH | | | ______ |
| 9979 DEBORAH KAY WAYNE | | Voted | ______ |
| 9979 JOEL R WAYNE | | Voted | ______ |

Figure 8: Neighbor mailer from Gerber, Green, and Larimer (2008). This mailer increased turnout rates by 8.1 percentage points, the largest effect that had ever been observed for a single-piece mailer.

> “With no other serious ‘suspects’ [sic] to investigate, it is difficult to make conclusive statements about Banksy’s identity based on the analysis presented here, other than saying the peaks of the geoprofiles in both Bristol and London include addresses known to be associated with [name redacted].”

Following Metcalf and Crawford (2016), I have decided to not to include the name of the individual when discussing this study.

a) Assess this study using the principles and ethical frameworks in this chapter.
b) Would you have done this study?
c) The authors justify this study in the abstract of the paper with the following sentence: “More broadly, these results support previous suggestions that the analysis of minor terrorism-related acts (e.g., graffiti) could be used to help locate terrorist bases before more serious incidents occur, and provides a fascinating example of the application of the model to a complex, real-world problem.” Does this change your opinion of the paper? If so, how?
d) The authors included the following ethical note at the end of their paper: “The authors are aware of, and respectful of, the privacy of [name redacted] and his relatives and have thus only used data in the public domain. We have deliberately omitted precise addresses.” Does this change your opinion of the paper? If so, how? Do you think the public/private dichotomy makes sense in this case?

10. [◠] In an interesting article Metcalf (2016) makes the argument that “publicly available datasets containing private data are among the most interesting to researchers and most risky to subjects.”

    a) What are two concrete examples that support this claim?
    b) In this same article Metcalf also claims that is anachronistic to assume that “any information harm has already been done by a public dataset”. Give one example of where this could be the case.

11. [◠] In this chapter I proposed the rule of thumb that *all* data is potentially identifiable and *all* data is potentially sensitive. Table 5 provides a list of examples of data that has no obviously personally identifying information but which can still be linked to specific people.

    a) Pick two of these examples and describe how the de-anonymization attack in both cases has a similar structure.
    b) For each of the two examples in part (a), describe how the data could reveal sensitive information about the people in the dataset.
    c) Now pick a third dataset from the table. Write an email to someone considering releasing it. Explain to them how this data could be potentially identifiable and potentially sensitive.

Table 5: List of examples of social data that does not have any obvious personally identifying information, but which can still be linked to specific people.

| Data | Citation |
|---|---|
| Health insurance records | Sweeney (2002) |
| Credit card transaction data | Montjoye et al. (2015) |
| Netflix movie rating data | Narayanan and Shmatikov (2008) |
| Phone call meta-data | Mayer, Mutchler, and Mitchell (2016) |
| Search log data | Barbaro and Zeller Jr (2006) |
| Demographic, administrative, and social data about students | Zimmer (2010) |

Figure 9: Street art by Banksy in Cheltenham, England. Photo by Brian Robert Marshall. Source: Wikimedia Commons.

12. [◠] Putting yourself in everyone's shoes includes your participants and the general public, not just your peers. This distinction is illustrated in the case of the Jewish Chronic Disease Hospital (Katz, Capron, and Glass 1972, Ch. 1; Lerner 2004; Arras 2008).

    Dr. Chester M. Southam was a distinguished physician and researcher at Sloan-Kettering Institute for Cancer Research and an Associate Professor of Medicine at the Cornell University Medical College. On July 16, 1963, Southam and two colleagues injected live cancer cells into the bodies of 22 debilitated patients at the Jewish Chronic Disease Hospital in New York. These injections were part of Southam's research to understand the immune system of cancerous patients. In earlier research, Southam had found that healthy volunteers were able to reject injected cancer cells in roughly 4 to 6 weeks, whereas it took patients who already had cancer much longer. Southam wondered whether the delayed response in the cancer patients was because they had cancer or because they were elderly and debilitated already. To address these possibilities, Southam decided to inject live cancer cells into a group of people who were elderly and debilitated but who did not have cancer. When word of the study spread, triggered in part by the resignation of three physicians who were asked to participate, some made comparisons to the Nazi Concentration Camp Experiments, but others—based in part on assurances by Southam—found the research unproblematic. Eventually, the New York State Board of Regents reviewed the case in order to decide if Southam should be able to continue to practice medicine. Southam argued at his defense that he was acting in "the best tradition of responsible clinical practice." Southam's defense was based on a number of claims, which were all supported by several distinguished experts who testified on his behalf: (1) his research was of high scientific and social merit; (2) there were no appreciable risks to participants; a claim based in part of Southam's 10 years of prior experience with more than 600 subjects; (3) the level of disclosure should be adjusted according to the level of risk posed by the researcher; (4) the research was in conformity with the standard of medical practice at that time. Ultimately, the Regent's board found Southam guilty of fraud, deceit, and unprofessional conduct, and suspended his medical license for one year. Yet, just a few years later, Chester M. Southam was elected president of the American Association of Cancer Researchers.

    a) Assess Southam's study using the four principles in this chapter.
    b) It appears that Southam took the perspective of his colleagues and correctly anticipated how they might respond to his work; in fact, many of them testified on his behalf. But, he was unable or unwilling to understand how his research might be troubling to the public. What role do you think public opinion—which could be distinct from the opinions of participants—should have in research ethics? What should happen if popular opinion and peer opinion differ?

13. [◠] In a paper titled "Crowdseeding in Eastern Congo: Using Cell Phones to Collect Conflict Events Data in Real Time", Van der Windt and Humphreys (2016) describe a distributed data collection system (see Chapter 5) that they created in Eastern Congo. Describe how the researchers dealt with the uncertainty about possible harms to participants.

14. [◠] In October 2014, three political scientists sent mailers to 102,780 registered voters in Montana as part of an experiment to measure whether voters who are given more information are more likely to vote. The mailers—which were labeled 2014 Montana General Election Voter Information Guide—placed Montana Supreme Court Justice candidates, which is a non-partisan election, on a scale from liberal to conservative, which included Barack Obama and Mitt Romney as comparisons. The mailer also included a reproduction of the Great Seal of the State of Montana (Figure 10).

    The mailers generated complaints from Montana voters, and they caused Linda McCulloch, Montana's Secretary of State, to file a formal complaint with the Montana state government. The universities that employed the researchers—Dartmouth and Stanford—sent a letter to everyone that had received the mailer apologizing for any potential confusion and making clear that the mailer "was not affiliated with any political party, candidate or organization, and was not intended to influence any race." The letter also clarified that the ranking "relied upon public information about who had donated to each of the campaigns." (Figure 11)

In May 2015, the Commissioner of Political Practices of the State of Montana, Jonathan Motl, determined that the researchers violated Montana law: "The Commissioner determines that there are sufficient facts to show that Stanford, Dartmouth and/or its researchers violated Montana campaign practice laws requiring registration, reporting and disclosure of independent expenditures." (Sufficient Finding Number 3 in Motl (2015)). The Commissioner also recommended that the County Attorney investigate whether the use of the unauthorized use of the Great Seal of Montana violates Montana state law (Motl 2015).

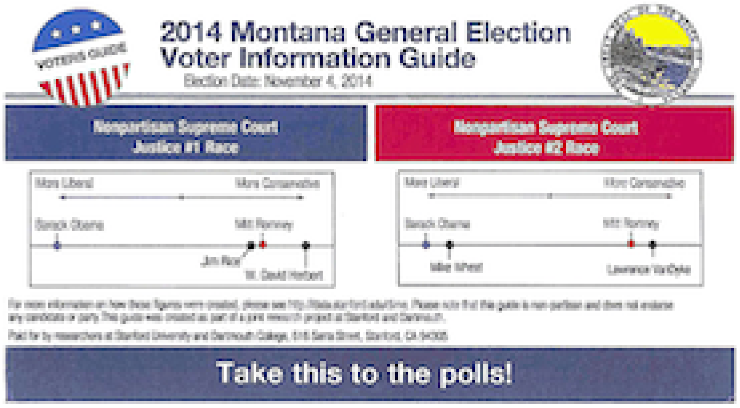
Stanford and Dartmouth disagreed with Motl's ruling. A Stanford spokeswoman named Lisa Lapin said "Stanford…does not believe any election laws were violated" and that the mailing "did not contain any advocacy supporting or opposing any candidate." She pointed out that the mailer explicitly stated that it "is nonpartisan and does not endorse any candidate or party." (Richman 2015)

a) Assess this study using the four principles and two frameworks described in this chapter.
b) Assume that the mailers were sent to a random sample of voters (but more on that in a moment), under what conditions might this mailing have altered the outcome of the Supreme Court Justice election?
c) In fact, the mailers were not sent to a random sample of voters. According to a report by Jeremy Johnson (a political scientists who assisted in the investigation), mailers "were sent to 64,265 voters identified as likely liberal to centrist leaning in Democratic leaning precincts and 39,515 voters identified as conservative to centrist in Republican leaning precincts. The researchers justified the disparity between Democratic and Republican numbers on grounds that they anticipated turnout to be significantly lower among Democratic voters." Does this change your assessment of the research design? If so, how?
d) In response to the investigation, the researchers said that they picked this election in part because "neither judicial race had been closely contested in the primary. Based on an analysis of the 2014 primary election results in the context of previous Montana judicial elections, the researchers determined that the research study as designed would not change the outcome of either contest." (Motl 2015) Does this change your assessment of the research? If so, how?
e) In fact, the election turned out to be not particularly close (Table 6). Does this change your assessment of the research? If so, how?
f) It turns out that a study was submitted to Dartmouth IRB by one of the researchers, but it differed substantially from the actual Montana study. The mailer used in Montana was never submitted to the IRB. The study was never submitted to the Stanford IRB. Does this change your assessment of the research? If so, how?
g) It also turns out that the researchers sent similar election materials to 143,000 voters in California and 66,000 in New Hampshire. As far as I know, there were no formal complaints triggered by these approximately 200,000 additional mailers. Does this change your assessment of the research? If so, how?
h) What, if anything, would you have done differently if you were the principal investigators? How would you have designed the study if you were interested in exploring whether additional information increases voter turnout in nonpartisan races?

Table 6: Results from the 2014 Montana Supreme Court Justice elections.
Source: Webpage of Montana Secretary of State.

| Candidates | Votes received | Percentage |
|---|---:|---:|
| Supreme Court Justice #1 | | |
| W. David Herbert | 65,404 | 21.59% |
| Jim Rice | 236,963 | 78.22% |

| Candidates | Votes received | Percentage |
|---|---|---|
| Supreme Court Justice #2 | | |
| Lawrence VanDyke | 134,904 | 40.80% |
| Mike Wheat | 195,303 | 59.06% |



2014 Montana General Election
Voter Information Guide
Election Date: November 4, 2014

Nonpartisan Supreme Court Justice #1 Race

Nonpartisan Supreme Court Justice #2 Race

Take this to the polls!

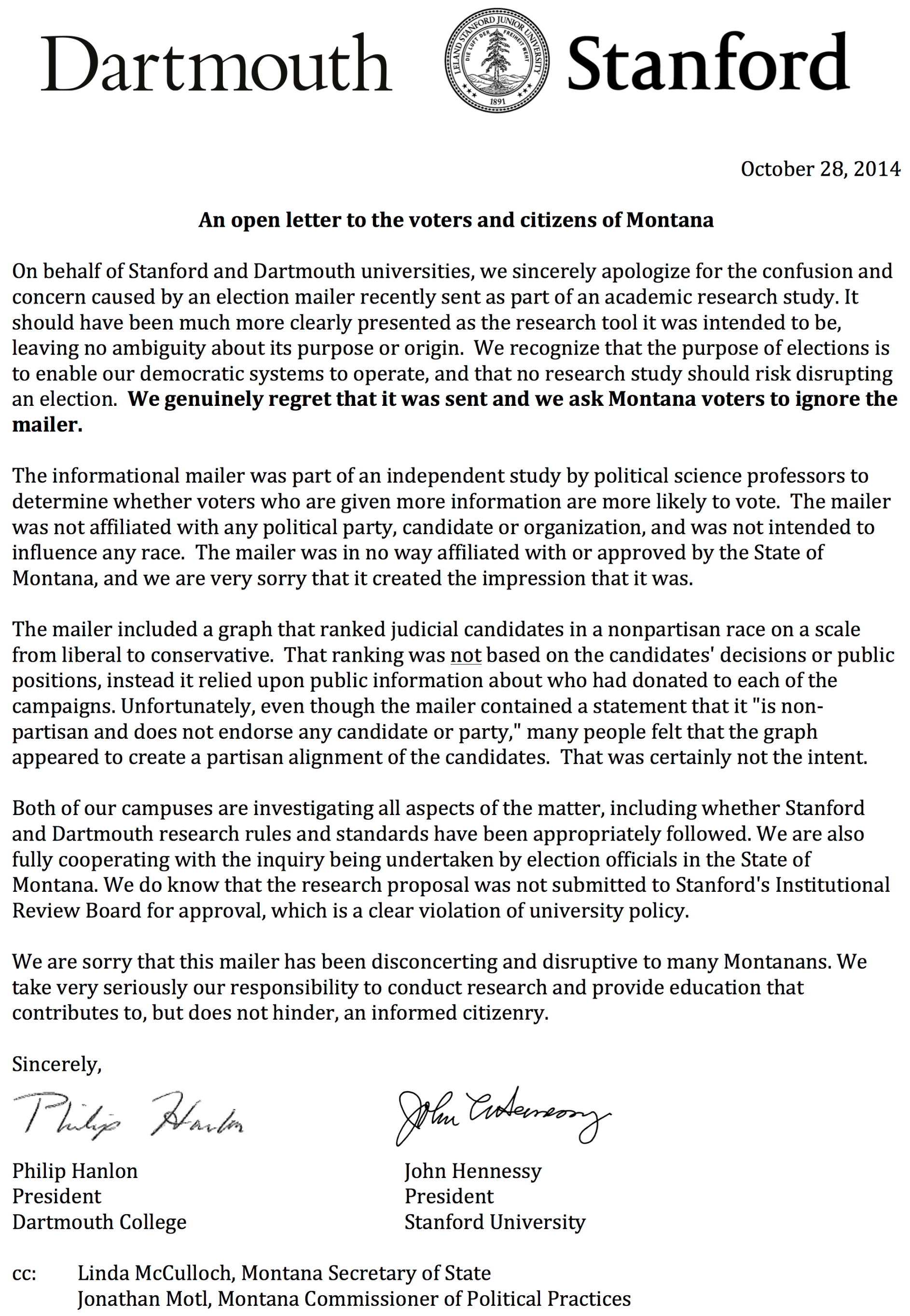
Figure 10: Mailer sent by three political scientists to 102,780 registered voters in Montana as part of an experiment to measure whether voters who are given more information are more likely to vote. The sample size in this experiment was roughly 15% of eligible voters in the state.

15. [◔] On May 8, 2016, two researchers—Emil Kirkegaard and Julius Bjerrekaer—scraped information from the online dating site OkCupid and publicly released a dataset of about 70,000 users, including variables of username, age, gender, location, religion-related opinions, astrology-related opinions, dating interests, number of photos, etc., as well as answers given to the top 2600 questions on the site. In a draft paper accompanying the released data, the authors stated that "Some may object to the ethics of gathering and releasing this data. However, all the data found in the dataset are or were already publicly available, so releasing this dataset merely presents it in a more useful form."

    In response to the data release, one of the authors was asked on Twitter: "This data set is highly re-identifiable. Even includes usernames? Was any work at all done to anonymize it?". His response was "No. Data is already public." (Zimmer 2016; Resnick 2016)

    a) Assess this data release using the principles and ethical frameworks discussed in this chapter.
    b) Would you use this data for your own research?
    c) What if you scraped it yourself?

16. [◔] In 2010 an intelligence analyst with the U.S. Army gave 250,000 classified diplomatic cables to the organization WikiLeaks, and they were subsequently posted online. Gill and Spirling (2015) argue that "the WikiLeaks disclosure potentially represents a trove of data that might be tapped to test subtle theories in international relations", and then statistically characterize the sample of leaked documents. For example, the authors estimate that they represent about 5% of all diplomatic cables during that time period, but that this proportion varies from embassy to embassy (see Figure 1 of their paper).



Dartmouth Stanford

October 28, 2014

**An open letter to the voters and citizens of Montana**

On behalf of Stanford and Dartmouth universities, we sincerely apologize for the confusion and concern caused by an election mailer recently sent as part of an academic research study. It should have been much more clearly presented as the research tool it was intended to be, leaving no ambiguity about its purpose or origin. We recognize that the purpose of elections is to enable our democratic systems to operate, and that no research study should risk disrupting an election. **We genuinely regret that it was sent and we ask Montana voters to ignore the mailer.**

The informational mailer was part of an independent study by political science professors to determine whether voters who are given more information are more likely to vote. The mailer was not affiliated with any political party, candidate or organization, and was not intended to influence any race. The mailer was in no way affiliated with or approved by the State of Montana, and we are very sorry that it created the impression that it was.

The mailer included a graph that ranked judicial candidates in a nonpartisan race on a scale from liberal to conservative. That ranking was <u>not</u> based on the candidates' decisions or public positions, instead it relied upon public information about who had donated to each of the campaigns. Unfortunately, even though the mailer contained a statement that it "is non-partisan and does not endorse any candidate or party," many people felt that the graph appeared to create a partisan alignment of the candidates. That was certainly not the intent.

Both of our campuses are investigating all aspects of the matter, including whether Stanford and Dartmouth research rules and standards have been appropriately followed. We are also fully cooperating with the inquiry being undertaken by election officials in the State of Montana. We do know that the research proposal was not submitted to Stanford's Institutional Review Board for approval, which is a clear violation of university policy.

We are sorry that this mailer has been disconcerting and disruptive to many Montanans. We take very seriously our responsibility to conduct research and provide education that contributes to, but does not hinder, an informed citizenry.

Sincerely,

Philip Hanlon
President
Dartmouth College

John Hennessy
President
Stanford University

cc: Linda McCulloch, Montana Secretary of State
Jonathan Motl, Montana Commissioner of Political Practices

Figure 11: Apology letter that was sent to the 102,780 registered voters in Montana who had received the mailer in Figure 10. The letter was sent by the Presidents of Dartmouth and Stanford, the universities that employed the researchers who sent the mailer.

  a) Read the paper, and then write an ethical appendix to it.
  b) The authors did not analyze the content of any of the leaked documents. Is there any project using these cables that you would conduct? Is there any project using these cables that you would not conduct?

17. [] In order to study how companies respond to complaints, a researcher sent fake complaint letters to 240 high-end restaurants in New York City. Here's an excerpt from the fictitious letter.

    "I am writing this letter to you because I am outraged about a recent experience I had at your restaurant. Not long ago, my wife and I celebrated our first anniversary. ... The evening became soured when the symptoms began to appear about four hours after eating. Extended nausea, vomiting, diarrhea, and abdominal cramps all pointed to one thing: food poisoning. It makes me furious just thinking that our special romantic evening became reduced to my wife watching me curl up in a fetal position on the tiled floor of our bathroom in between rounds of throwing up. ... Although it is not my intention to file any reports with the Better Business Bureau or the Department of Health, I want you, [name of the restaurateur], to understand what I went through in anticipation that you will respond accordingly."

    a) Evaluate this study using the principles and ethical frameworks described in this chapter. Given your assessment, would you do the study?
    b) Here's how the restaurants who received the letter reacted: "It was culinary chaos as owners, managers and chefs searched through computers for [name redacted] reservations or credit card records, reviewed menus and produce deliveries for possibly spoiled food, and questioned kitchen workers about possible lapses, all spurred by what both the university and the professor now concede was the business school study from hell." (Kifner 2001) Does this information change how you assess the study?
    c) As far as I know, this study was not reviewed by an IRB or any other third-party. Does that change how you assess the study? Why or why not?

18. [] Building on this previous question, I'd like you to compare this study to a completely different study that also involved restaurants. In this other study, Neumark and colleagues (1996) sent two male and two female college students with fabricated resumes to apply for jobs as waiters and waitresses at 65 restaurants in Philadelphia, in order to investigate sex discrimination in restaurant hiring. The 130 applications led to 54 interviews and 39 job offers. The study found statistically significant evidence of sex discrimination against women in high-price restaurants.

    a) Write an ethical appendix for study.
    b) Do you think this study is ethically different from the one described in the previous question. If so, how?

19. [] Some time around 2010, 6,548 professors in the United States received emails similar to this one.

    > "Dear Professor Salganik,
    >
    > I am writing you because I am a prospective Ph.D. student with considerable interest in your research. My plan is to apply to Ph.D. programs this coming fall, and I am eager to learn as much as I can about research opportunities in the meantime.
    >
    > I will be on campus today, and although I know it is short notice, I was wondering if you might have 10 minutes when you would be willing to meet with me to briefly talk about your work and any possible opportunities for me to get involved in your research. Any time that would be convenient for you would be fine with me, as meeting with you is my first priority during this campus visit.

> Thank you in advance for your consideration.
>
> Sincerely, Carlos Lopez"

These emails were part of a field experiment to measure whether professors were more likely to respond to the email depending on 1) the time-frame (today vs next week) and 2) the name of the sender which was varied to signal ethnicity and gender (e.g., Meredith Roberts, Raj Singh, etc). The researchers found that when the requests were to meet in 1 week, Caucasian males were granted access to faculty members about 25% more often than were women and minorities. But, when the fictitious students requested meetings that same day these patterns were essentially eliminated (Milkman, Akinola, and Chugh 2012).

a) Assess this experiment according to the principles and frameworks in this chapter.
b) After the study was over, the researchers sent the following debriefing email to all participants.

> "Recently, you received an email from a student asking for 10 minutes of your time to discuss your Ph.D. program (the body of the email appears below). We are emailing you today to debrief you on the actual purpose of that email, as it was part of a research study. We sincerely hope our study did not cause you any disruption and we apologize if you were at all inconvenienced. Our hope is that this letter will provide a sufficient explanation of the purpose and design of our study to alleviate any concerns you may have about your involvement. We want to thank you for your time and for reading further if you are interested in understanding why you received this message. We hope you will see the value of the knowledge we anticipate producing with this large academic study."

After explaining the purpose and design of the study, they further noted that:

> "As soon as the results of our research are available, we will post them on our websites. Please rest assured that no identifiable data will ever be reported from this study, and our between subject design ensures that we will only be able to identify email responsiveness patterns in aggregate – not at the individual level. No individual or university will be identifiable in any of the research or data we publish. Of course, any one individual email response is not meaningful as there are multiple reasons why an individual faculty member might accept or decline a meeting request. All data has already been de-identified and the identifiable email responses have already been deleted from our databases and related server. In addition, during the time when the data was identifiable, it was protected with strong and secure passwords. And as is always the case when academics conduct research involving human subjects, our research protocols were approved by our universities' Institutional Review Boards (the Columbia University Morningside IRB and the University of Pennsylvania IRB).
>
> If you have any questions about your rights as a research subject, you may contact the Columbia University Morningside Institutional Review Board at 212-851-7040 or by email at askirb@columbia.edu and/or the University of Pennsylvania Institutional Review Board at 215-898-2614.
>
> Thank you again for your time and understanding of the work we are doing."

c) What are the arguments for debriefing in this case? What are the arguments against? Do you think that the researchers should have debriefed the participants in this case?
d) In the supporting online materials, the researchers have a section titled "Human Subjects Protections." Please read this section. Is there anything that you would add or remove.

e) What was the cost of this experiment to the researchers? What was the cost of this experiment to participants? Andrew Gelman (2010) has argued that participants in this study could have been compensated for their time after the experiments was over. Do you agree? Try to make your argument using the principles and ethical frameworks in the chapter.

## References

Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. 2015. "Privacy and Human Behavior in the Age of Information." *Science* 347 (6221): 509–14. doi:10.1126/science.aaa1465.

Alexander, Larry, and Michael Moore. 2015. "Deontological Ethics." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2015. http://plato.stanford.edu/archives/spr2015/entries/ethics-deontological/.

Alsan, Marcella, and Marianne Wanamaker. 2016. "Tuskegee and the Health of Black Men." Working Paper 22323. National Bureau of Economic Research. http://www.nber.org/papers/w22323.

Angiuli, Olivia, Joe Blitzstein, and Jim Waldo. 2015. "How to de-Identify Your Data." *Communications of the ACM* 58 (12): 48–55. doi:10.1145/2814340.

Arras, John D. 2008. "The Jewish Chronic Disease Hospital Case." In *The Oxford Textbook of Clinical Research Ethics*, edited by E. J. Emanuel, R. A. Crouch, C. Grady, R. K. Lie, F. G. Miller, and D. Wendler, 73–79.

Baele, Stéphane J. 2013. "The Ethics of New Development Economics: Is the Experimental Approach to Development Economics Morally Wrong?" https://ore.exeter.ac.uk/repository/handle/10871/17048.

Bailey, Michael, David Dittrich, and Erin Kenneally. 2013. "Applying Ethical Principles to Information and Communications Technology Research: A Companion to the Menlo Report." Washington, DC: Department of Homeland Security.

Bankston, Kevin S., and Ashkan Soltani. 2013. "Tiny Constables and the Cost of Surveillance: Making Cents Out of United States V. Jones." *Yale LJF* 123: 335. http://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/yljfor123&section=14.

Barbaro, Michael, and Tom Zeller Jr. 2006. "A Face Is Exposed for AOL Searcher No. 4417749." *The New York Times*, August. http://select.nytimes.com/gst/abstract.html?res=F10612FC345B0C7A8CDDA10894DE404482.

Batson, C. Daniel, Shannon Early, and Giovanni Salvarani. 1997. "Perspective Taking: Imagining How Another Feels Versus Imaging How You Would Feel." *Personality and Social Psychology Bulletin* 23 (7): 751–58. doi:10.1177/0146167297237008.

Baumrind, Diana. 1985. "Research Using Intentional Deception: Ethical Issues Revisited." *American Psychologist* 40 (2): 165–74. doi:10.1037/0003-066X.40.2.165.

Beauchamp, Tom L. 2011. "The Belmont Report." In *The Oxford Textbook of Clinical Research Ethics*, edited by Ezekiel J. Emanuel, Christine C. Grady, Robert A. Crouch, Reidar K. Lie, Franklin G. Miller, and David D. Wendler. Oxford University Press.

Beauchamp, Tom L., and James F. Childress. 2012. *Principles of Biomedical Ethics*. 7 edition. New York: Oxford University Press.

Beauchamp, Tom L., and Yashar Saghai. 2012. "The Historical Foundations of the Research-Practice

Distinction in Bioethics." *Theoretical Medicine and Bioethics* 33 (1): 45–56. doi:10.1007/s11017-011-9207-8.

Bederson, Benjamin B., and Alexander J. Quinn. 2011. "Web Workers Unite! Addressing Challenges of Online Laborers." In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, 97–106. CHI EA '11. New York, NY, USA: ACM. doi:10.1145/1979742.1979606.

Bell, Gordon, and Jim Gemmell. 2009. *Total Recall: How the E-Memory Revolution Will Change Everything.* First Edition edition. New York: Dutton Adult.

Belmont Report. 1979. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research.* US Deptartment of Health, Education, and Welfare.

Bengtsson, Linus, Xin Lu, Anna Thorson, Richard Garfield, and Johan von Schreeb. 2011. "Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in {Haiti}." *PLoS Med* 8 (8): e1001083. doi:10.1371/journal.pmed.1001083.

Blumenstock, Joshua E., Gabriel Cadamuro, and Robert On. 2015. "Predicting Poverty and Wealth from Mobile Phone Metadata." *Science* 350 (6264): 1073–6. doi:10.1126/science.aac4420.

boyd, danah. 2016. "Untangling Research and Practice: What Facebooks Emotional Contagion Study Teaches Us." *Research Ethics* 12 (1): 4–13. doi:10.1177/1747016115583379.

boyd, danah, and Kate Crawford. 2012. "Critical Questions for Big Data." *Information, Communication & Society* 15 (5): 662–79. doi:10.1080/1369118X.2012.678878.

Brandt, Allan M. 1978. "Racism and Research: The Case of the Tuskegee Syphilis Study." *The Hastings Center Report* 8 (6): 21–29. doi:10.2307/3561468.

Brickell, Justin, and Vitaly Shmatikov. 2008. "The Cost of Privacy: Destruction of Data-Mining Utility in Anonymized Data Publishing." In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 70–78. KDD '08. doi:10.1145/1401890.1401904.

Brownstein, John S., Christopher A. Cassa, and Kenneth D. Mandl. 2006. "No Place to Hide Reverse Identification of Patients from Published Maps." *New England Journal of Medicine* 355 (16): 1741–2. doi:10.1056/NEJMc061891.

Bruckman, Amy. 2016a. "Do Researchers Need to Abide by Terms of Service (TOS)? An Answer." *The Next Bison: Social Computing and Culture.* https://nextbison.wordpress.com/2016/02/26/tos/.

———. 2016b. "More on TOS: Maybe Documenting Intent Is Not So Smart." *The Next Bison: Social Computing and Culture.* https://nextbison.wordpress.com/2016/02/29/tos2/.

Burnett, Sam, and Nick Feamster. 2015. "Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests." In *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, 653–67. SIGCOMM '15. London, England: ACM. doi:10.1145/2785956.2787485.

Calo, Ryan. 2011. "The Boundaries of Privacy Harm." *Indiana Law Journal* 86: 1131. http://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/indana86&section=30.

———. 2013. "Consumer Subject Review Boards: A Thought Experiment." *Stanford Law Review Online*, 97–102. http://papers.ssrn.com/abstract=2340745.

Carter, Stephen L. 1996. *Integrity.* Harper Collins.

Chaabane, Abdelberi, Terence Chen, Mathieu Cunche, Emiliano De Cristofaro, Arik Friedman, and Mohamed Ali Kaafar. 2014. "Censorship in the Wild: Analyzing Internet Filtering in Syria." In

*Proceedings of the 2014 Conference on Internet Measurement Conference*, 285–98. IMC '14. New York, NY, USA: ACM. doi:10.1145/2663716.2663720.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences.* Second edition. Hillsdale, N.J: Routledge.

Cornwell, Erin York. 2010. "Opening and Closing the Jury Room Door: A Sociohistorical Consideration of the 1955 Chicago Jury Project Scandal." *Justice System Journal* 31 (1): 49–73. doi:10.1080/0098261X.2010.10767954.

Council, National Research. 2014. *Proposed Revisions to the Common Rule for the Protection of Human Subjects in the Behavioral and Social Sciences.* Committee on Revisions to the Common Rule for the Protection of Human Subjects in Research in the Behavioral and Social Sciences. Board on Behavioral, Cognitive, and Sensory Sciences, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

Crawford, Kate. 2014. "The Test We Canand ShouldRun on Facebook." *The Atlantic*, July. http://www.theatlantic.com/technology/archive/2014/07/the-test-we-canand-shouldrun-on-facebook/373819/.

Currie, Janet. 2013. "Big Data Versus Big Brother: On the Appropriate Use of Large-Scale Data Collections in Pediatrics." *Pediatrics* 131 (Supplement 2): S127–S132. doi:10.1542/peds.2013-0252c.

Daries, Jon P., Justin Reich, Jim Waldo, Elise M. Young, Jonathan Whittinghill, Daniel Thomas Seaton, Andrew Dean Ho, and Isaac Chuang. 2014. "Privacy, Anonymity, and Big Data in the Social Sciences." *Queue* 12 (7): 30:30–30:41. doi:10.1145/2639988.2661641.

Desai, Tanvi, Felix Ritchie, and Richard Welpton. 2016. "Five Safes: Designing Data Access for Research." University of the West of England. http://www1.uwe.ac.uk/bl/research/bristoleconomicanalysis/economicsworkingpapers/economicspapers2016.aspx.

Desposato, Scott. 2016a. "Conclusion and Recommendations." In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, edited by Scott Desposato. Routledge.

———. 2016b. *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals.* Routledge.

Dickert, Neal, and Christine Grady. 2008. "Incentives for Research Participants." In *The Oxford Textbook of Clinical Research Ethics*, edited by E. J. Emanuel, R. A. Crouch, C. Grady, R. K. Lie, F. G. Miller, and D. Wendler, 386–97.

Dittrich, D., K. Carpenter, and M. Karir. 2015. "The Internet Census 2012 Dataset: An Ethical Examination." *IEEE Technology and Society Magazine* 34 (2): 40–46. doi:10.1109/MTS.2015.2425592.

Dittrich, David, Erin Kenneally, and others. 2011. "The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research." *US Department of Homeland Security.* http://www.caida.org/publications/papers/2012/menlo_report_ethical_principles/.

Dwork, Cynthia, and Aaron Roth. 2014. "The Algorithmic Foundations of Differential Privacy." *Found. Trends Theor. Comput. Sci.* 9 (34): 211–407. doi:10.1561/0400000042.

Dwork, Cynthia, Adam D. Smith, Thomas Steinke, and Jonathan Ullman. 2017. "Hiding in Plain Sight: A Survey of Attacks on Private Data." *Annual Review of Statistics and Its Application* 4 (1): null. doi:10.1146/annurev-statistics-060116-054123.

Emanuel, Ezekiel J., Christine C. Grady, Robert A. Crouch, Reidar K. Lie, Franklin G. Miller, and David D. Wendler, eds. 2008. *The Oxford Textbook of Clinical Research Ethics.* Oxford University

Press.

Emanuel, Ezekiel J., David Wendler, and Christine Grady. 2000. "What Makes Clinical Research Ethical?" *Journal of the American Medical Association* 283 (20): 2701–11. doi:10.1001/jama.283.20.2701.

Engineering, Institute of Medicine and National Academy of Sciences and National Academy of. 2009. *On Being a Scientist: A Guide to Responsible Conduct in Research.* Third. Washington, DC: National Academies Press. dx.doi.org/10.17226/12192.

Enserink, Martin, and Gilbert Chin. 2015. "The End of Privacy." *Science* 347 (6221): 490–91. doi:10.1126/science.347.6221.490.

Evans, Barbara J. 2013. "Why the Common Rule Is Hard to Amend." SSRN Scholarly Paper ID 2183701. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2183701.

Eyal, Nir. 2012. "Informed Consent." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2012. http://plato.stanford.edu/archives/fall2012/entries/informed-consent/.

Farber, Henry S. 2015. "Why You Cant Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers." *The Quarterly Journal of Economics* 130 (4): 1975–2026. doi:10.1093/qje/qjv026.

Finn, P., and M. Jakobsson. 2007. "Designing Ethical Phishing Experiments." *IEEE Technology and Society Magazine* 26 (1): 46–58. doi:10.1109/MTAS.2007.335565.

Fiore-Gartland, Brittany, and Gina Neff. 2015. "Communication, Mediation, and the Expectations of Data: Data Valences Across Health and Wellness Communities." *International Journal of Communication* 9: 19. http://ijoc.org/index.php/ijoc/article/view/2830.

Fiske, Susan T., and Robert M. Hauser. 2014. "Protecting Human Research Participants in the Age of Big Data." *Proceedings of the National Academy of Sciences* 111 (38): 13675–6. doi:10.1073/pnas.1414626111.

Flick, Catherine. 2015. "Informed Consent and the Facebook Emotional Manipulation Study." *Research Ethics*, August, 1747016115599568. doi:10.1177/1747016115599568.

Foucault, Michel. 1995. *Discipline & Punish: The Birth of the Prison.* Translated by Alan Sheridan. New York: Vintage Books.

Gelman, Andrew. 2010. "63,000 Worth of Abusive Research . . . or Just a Really Stupid Waste of Time?" *Statistical Modeling, Causal Inference, and Social Science.* http://andrewgelman.com/2010/05/06/63000_worth_of/.

Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 102 (01): 33–48. doi:10.1017/S000305540808009X.

Gill, Michael, and Arthur Spirling. 2015. "Estimating the Severity of the WikiLeaks U.S. Diplomatic Cables Disclosure." *Political Analysis* 23 (2): 299–305. doi:10.1093/pan/mpv005.

Gillon, Raanan. 2015. "Defending the Four Principles Approach as a Good Basis for Good Medical Practice and Therefore for Good Medical Ethics." *Journal of Medical Ethics* 41 (1): 111–16. doi:10.1136/medethics-2014-102282.

Goroff, Daniel L. 2015. "Balancing Privacy Versus Accuracy in Research Protocols." *Science* 347 (6221): 479–80. doi:10.1126/science.aaa3483.

Grimmelmann, James. 2015. "The Law and Ethics of Experiments on Social Media Users." SSRN

Scholarly Paper ID 2604168. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2604776.

Hauge, Michelle V., Mark D. Stevenson, D. Kim Rossmo, and Steven C. Le Comber. 2016. "Tagging Banksy: Using Geographic Profiling to Investigate a Modern Art Mystery." *Journal of Spatial Science* 61 (1): 185–90. doi:10.1080/14498596.2016.1138246.

Heller, Jean. 1972. "Syphilis Victims in U.S. Study Went Untreated for 40 Years." *New York Times*, July, 1 and 8.

Hernandez, Daniela, and Deepa Seetharaman. 2016. "Facebook Offers Details on How It Handles Research." *Wall Street Journal*, June. http://www.wsj.com/articles/facebook-offers-details-how-it-handles-research-14659

Holm, S. 1995. "Not Just Autonomy–the Principles of American Biomedical Ethics." *Journal of Medical Ethics* 21 (6): 332–38. doi:10.1136/jme.21.6.332.

Holmes, David S. 1976a. "Debriefing After Psychological Experiments: I. Effectiveness of Postdeception Dehoaxing." *American Psychologist* 31 (12): 858–67. doi:10.1037/0003-066X.31.12.858.

———. 1976b. "Debriefing After Psychological Experiments: II. Effectiveness of Postexperimental Desensitizing." *American Psychologist* 31 (12): 868–75. doi:10.1037/0003-066X.31.12.868.

Hoonaard, Will C. van den. 2011. *Seduction of Ethics: Transforming the Social Sciences*. Toronto ; Buffalo N.Y.: University of Toronto Press, Scholarly Publishing Division.

Hudson, Kathy L., and Francis S. Collins. 2015. "Bringing the Common Rule into the 21st Century." *New England Journal of Medicine* 373 (24): 2293–6. doi:10.1056/NEJMp1512205.

Humphreys, Macartan. 2015. "Reflections on the Ethics of Social Experimentation." *Journal of Globalization and Development* 6 (1): 87–112. doi:10.1515/jgd-2014-0016.

Hunter, David, and Nicholas Evans. 2016. "Facebook Emotional Contagion Experiment Controversy." *Research Ethics* 12 (1): 2–3. doi:10.1177/1747016115626341.

Issenberg, Sasha. 2012. *The Victory Lab: The Secret Science of Winning Campaigns*. New York: Broadway Books.

Jackman, Molly, and Lauri Kanerva. 2016. "Evolving the IRB: Building Robust Review for Industry Research." *Washington and Lee Law Review Online* 72 (3): 442. http://scholarlycommons.law.wlu.edu/wlulr-online/vol72/iss3/8.

Jagatic, Tom N., Nathaniel A. Johnson, Markus Jakobsson, and Filippo Menczer. 2007. "Social Phishing." *Commun. ACM* 50 (10): 94–100. doi:10.1145/1290958.1290968.

Jakobsson, Markus, and Jacob Ratkiewicz. 2006. "Designing Ethical Phishing Experiments: A Study of (ROT13) rOnl Query Features." In *Proceedings of the 15th International Conference on World Wide Web*, 513–22. WWW '06. New York, NY, USA: ACM. doi:10.1145/1135777.1135853.

Jones, Ben, and Nick Feamster. 2015. "Can Censorship Measurements Be Safe(R)?" In *Proceedings of the 14th ACM Workshop on Hot Topics in Networks*, 1:1–1:7. HotNets-XIV. New York, NY, USA: ACM. doi:10.1145/2834050.2834066.

Jones, James H. 1993. *Bad Blood: The Tuskegee Syphilis Experiment, New and Expanded Edition*. Revised edition. New York : Toronto : New York: Free Press.

———. 2011. "The Tuskegee Syphilis Experiment." In *The Oxford Textbook of Clinical Research Ethics*, edited by Ezekiel J. Emanuel, Christine C. Grady, Robert A. Crouch, Reidar K. Lie, Franklin G. Miller, and David D. Wendler. Oxford University Press.

Kahn, Jeffrey P., Effy Vayena, and Anna C. Mastroianni. 2014. "Opinion: Learning as We Go:

Lessons from the Publication of Facebooks Social-Computing Research." *Proceedings of the National Academy of Sciences* 111 (38): 13677–9. doi:10.1073/pnas.1416405111.

Katz, Jay, Alexander Morgan Capron, and Eleanor Swift Glass. 1972. *Experimentation with Human Beings: The Authority of the Investigator, Subject, Professions, and State in the Human Experimentation Process.* Russell Sage Foundation. http://www.jstor.org/stable/10.7758/9781610448345.

Kifner, John. 2001. "Scholar Sets Off Gastronomic False Alarm." *The New York Times*, September. http://www.nytimes.com/2001/09/08/nyregion/scholar-sets-off-gastronomic-false-alarm.html.

King, Gary, and Melissa Sands. 2015. "How Human Subjects Research Rules Mislead You and Your University, and What to Do About It." *Working Paper*, August. http://j.mp/1d2gSQQ.

Kleinsman, John, and Sue Buckley. 2015. "Facebook Study: A Little Bit Unethical But Worth It?" *Journal of Bioethical Inquiry* 12 (2): 179–82. doi:10.1007/s11673-015-9621-0.

Klitzman, Robert. 2015. *The Ethics Police?: The Struggle to Make Human Research Safe.* Oxford ; New York: Oxford University Press.

Kosinski, Michal, Sandra C. Matz, Samuel D. Gosling, Vesselin Popov, and David Stillwell. 2015. "Facebook as a Research Tool for the Social Sciences: Opportunities, Challenges, Ethical Considerations, and Practical Guidelines." *American Psychologist* 70 (6): 543–56. doi:10.1037/a0039210.

Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior." *Proceedings of the National Academy of Sciences*, March. doi:10.1073/pnas.1218772110.

Kramer, Adam D. I., Jamie E. Guillory, and Jeffrey T. Hancock. 2014. "Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks." *Proceedings of the National Academy of Sciences* 111 (24): 8788–90. doi:10.1073/pnas.1320040111.

Landau, Susan. 2016. "Transactional Information Is Remarkably Revelatory." *Proceedings of the National Academy of Sciences* 113 (20): 5467–9. doi:10.1073/pnas.1605356113.

Lanier, Jaron. 2014. *Who Owns the Future?* Reprint edition. New York: Simon & Schuster.

Lerner, Barron H. 2004. "Sins of Omission Cancer Research Without Informed Consent." *New England Journal of Medicine* 351 (7): 628–30. doi:10.1056/NEJMp048108.

Lewis, Kevin, Marco Gonzalez, and Jason Kaufman. 2012. "Social Selection and Peer Influence in an Online Social Network." *Proceedings of the National Academy of Sciences* 109 (1): 68–72. doi:10.1073/pnas.1109739109.

Lewis, Kevin, Jason Kaufman, Marco Gonzalez, Andreas Wimmer, and Nicholas Christakis. 2008. "Tastes, Ties, and Time: A New Social Network Dataset Using Facebook.com." *Social Networks* 30 (4): 330–42. doi:10.1016/j.socnet.2008.07.002.

Lowrance, William W. 2012. *Privacy, Confidentiality, and Health Research.* Cambridge University Press.

Lu, Xin, Linus Bengtsson, and Petter Holme. 2012. "Predictability of Population Displacement After the 2010 Haiti Earthquake." *Proceedings of the National Academy of Sciences* 109 (29): 11576–81. doi:10.1073/pnas.1203882109.

MacCarthy, Mark. 2015. "Privacy Restrictions and Contextual Harm." *Working Paper.*

Maddock, Jim, Robert Mason, and Kate Starbird. 2015. "Using Historical Twitter Data for Research: Ethical Challenges of Tweet Deletions." In *CSCW 15 Workshop on Ethics at the 2015 Conference* *on Computer Supported Cooperative Work*. Vancouver, Canada.

Manson, Neil C., and Onora O'Neill. 2007. *Rethinking Informed Consent in Bioethics*. Cambridge University Press.

Markham, Annette, and Elizabeth Buchanan. 2012. "Ethical Decision-Making and Internet Research: Recommendations from the AoIR Ethics Working Committee." Version 2.0. Association of Internet Researchers. https://cms.bsu.edu/sitecore/shell/-/media/WWW/DepartmentalContent/ResearchIntegrity/Files/Education/Active/AoIR%20Social%20Media%20Working%20Committee.pdf.

Marx, Gary T. 2016. *Windows Into the Soul: Surveillance and Society in an Age of High Technology*. University of Chicago Press.

Mastroianni, Anna, and Jeffrey Kahn. 2001. "Swinging on the Pendulum: Shifting Views of Justice in Human Subjects Research." *Hastings Center Report* 31 (3): 21–28. doi:10.2307/3527551.

Mayer, Jonathan, Patrick Mutchler, and John C. Mitchell. 2016. "Evaluating the Privacy Properties of Telephone Metadata." *Proceedings of the National Academy of Sciences* 113 (20): 5536–41. doi:10.1073/pnas.1508081113.

Mayer-Schönberger, Viktor. 2009. *Delete: The Virtue of Forgetting in the Digital Age*. Princeton University Press.

McDonald, Sean. 2016. "Ebola: A Big Data Disaster." CIS Papers 2016.01. The Centre for Internet & Society. http://cis-india.org/papers/ebola-a-big-data-disaster.

Metcalf, Jacob. 2016. "Big Data Analytics and Revision of the Common Rule." *Communications of the ACM* 59 (7): 31–33. doi:10.1145/2935882.

Metcalf, Jacob, and Kate Crawford. 2016. "Where Are Human Subjects in Big Data Research? The Emerging Ethics Divide." *Big Data & Society* 3 (1): 1–14. doi:10.1177/2053951716650211.

Meyer, Michelle N. 2013. "Three Challenges for Risk-Based (Research) Regulation: Heterogeneity Among Regulated Activities, Regulator Bias, and Stakeholder Heterogeneity." SSRN Scholarly Paper ID 2218549. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2218549.

———. 2014. "Misjudgements Will Drive Social Trials Underground." *Nature* 511 (7509): 265–65. doi:10.1038/511265a.

———. 2015. "Two Cheers for Corporate Experimentation: The A/B Illusion and the Virtues of Data-Driven Innovation." SSRN Scholarly Paper ID 2605132. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2605132.

Middlemist, R. D., E. S. Knowles, and C. F. Matter. 1976. "Personal Space Invasions in the Lavatory: Suggestive Evidence for Arousal." *Journal of Personality and Social Psychology* 33 (5): 541–46.

Milkman, Katherine L., Modupe Akinola, and Dolly Chugh. 2012. "Temporal Distance and Discrimination An Audit Study in Academia." *Psychological Science* 23 (7): 710–17. doi:10.1177/0956797611434539.

Miller, Franklin G. 2014. "Clinical Research Before Informed Consent." *Kennedy Institute of Ethics Journal* 24 (2): 141–57.

Mills, Judson. 1976. "A Procedure for Explaining Experiments Involving Deception." *Personality and Social Psychology Bulletin* 2 (1): 3–13. doi:10.1177/014616727600200102.

Molloy, Jennifer C. 2011. "The Open Knowledge Foundation: Open Data Means Better Science."

*PLoS Biol* 9 (12): e1001195. doi:10.1371/journal.pbio.1001195.

Montjoye, Yves-Alexandre de, Laura Radaelli, Vivek Kumar Singh, and Alex Sandy Pentland. 2015. "Unique in the Shopping Mall: On the Reidentifiability of Credit Card Metadata." *Science* 347 (6221): 536–39. doi:10.1126/science.1256297.

Motl, Jonathan R. 2015. "McCulloch V. Stanford and Dartmouth." COPP 2014-CFP-046. Helena, Montana: Commisioner of Political Practices of the State of Montana. http://politicalpractices.mt.gov/content/2recentdecisions/McCullochvStanfordandDartmouthFinalDecision.

Narayanan, Arvind, and Vitaly Shmatikov. 2008. "Robust De-Anonymization of Large Sparse Datasets." In *Proceedings of the 2008 IEEE Symposium on Security and Privacy*, 111–25. Washington, DC, USA: IEEE Computer Society. doi:10.1109/SP.2008.33.

———. 2010. "Myths and Fallacies of 'Personally Identifiable Information'." *Commun. ACM* 53 (6): 24–26. doi:10.1145/1743546.1743558.

Narayanan, Arvind, and Bendert Zevenbergen. 2015. "No Encore for Encore? Ethical Questions for Web-Based Censorship Measurement." *Technology Science*, December. http://techscience.org/a/2015121501/.

Narayanan, Arvind, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder. 2016. *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction.* Princeton: Princeton University Press.

Narayanan, Arvind, Joanna Huey, and Edward W. Felten. 2016. "A Precautionary Approach to Big Data Privacy." In *Data Protection on the Move*, edited by Serge Gutwirth, Ronald Leenes, and Paul De Hert, 357–85. Law, Governance and Technology Series 24. Springer Netherlands. http://link.springer.com/chapter/10.1007/978-94-017-7376-8_13.

Neumark, David, Roy J. Bank, and Kyle D. Van Nort. 1996. "Sex Discrimination in Restaurant Hiring: An Audit Study." *The Quarterly Journal of Economics* 111 (3): 915–41. doi:10.2307/2946676.

Nickerson, David W., and Susan D. Hyde. 2016. "Conducting Research with NGOs: Relevant Counterfactuals from the Perspective of Subjects." In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, edited by Scott Desposato. Routledge.

Nissenbaum, Helen. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford Law Books.

———. 2011. "A Contextual Approach to Privacy Online." *Daedalus* 140 (4): 32–48. doi:10.1162/DAED_a_00113.

———. 2015. "Respecting Context to Protect Privacy: Why Meaning Matters." *Science and Engineering Ethics*, July. doi:10.1007/s11948-015-9674-9.

Oczak, Malgorzata, and Agnieszka Niedźwieńska. 2007. "Debriefing in Deceptive Research: A Proposed New Procedure." *Journal of Empirical Research on Human Research Ethics* 2 (3): 49–59. doi:10.1525/jer.2007.2.3.49.

Ohm, Paul. 2010. "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization." *UCLA Law Review* 57: 1701–77. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006.

———. 2015. "Sensitive Information." *Southern California Law Review* 88: 1125–96.

Packard, Vance. 1964. *The Naked Society.* D. McKay Company.

Pager, Devah. 2007. "The Use of Field Experiments for Studies of Employment Discrimination:

Contributions, Critiques, and Directions for the Future." *The Annals of the American Academy of Political and Social Science* 609: 104–33. http://www.jstor.org/stable/25097877.

Paluck, Elizabeth Levy, and Donald P. Green. 2009. "Deference, Dissent, and Dispute Resolution: An Experimental Intervention Using Mass Media to Change Norms and Behavior in Rwanda." *American Political Science Review* 103 (04): 622–44. doi:10.1017/S0003055409990128.

Panagopoulos, Costas. 2010. "Affect, Social Pressure and Prosocial Motivation: Field Experimental Evidence of the Mobilizing Effects of Pride, Shame and Publicizing Voting Behavior." *Political Behavior* 32 (3): 369–86. doi:10.1007/s11109-010-9114-0.

Parry, Marc. 2011. "Harvard Researchers Accused of Breaching Students' Privacy." *The Chronicle of Higher Education*, July. http://chronicle.com/article/Harvards-Privacy-Meltdown/128166/.

Penney, Jonathon. 2016. "Chilling Effects: Online Surveillance and Wikipedia Use." *Berkeley Technology Law Journal* 31 (1): 117. doi:http://dx.doi.org/10.15779/Z38SS13.

Plott, Charles R. 2013. "Risks and Harm: The Disciplines of Economics, Political Science, Game Theory, Judgement and Decision." National Academy of Sciences. http://www.tvworldwide.com/events/nas/130321/slides/Common%20Rule%20Workshop_Plott.pdf.

Polonetsky, Jules, Omer Tene, and Joseph Jerome. 2015. "Beyond the Common Rule: Ethical Structures for Data Research in Non-Academic Settings." SSRN Scholarly Paper ID 2621559. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2621559.

Porter, Joan P., and Greg Koski. 2008. "Regulations for the Protection of Humans in Research in the United States: The Common Rule." In *The Oxford Textbook of Clinical Research Ethics*. Oxford University Press.

Prentice, Deborah A., and Dale T. Miller. 1992. "When Small Effects Are Impressive." *Psychological Bulletin* 112 (1): 160–64. doi:10.1037/0033-2909.112.1.160.

Puschmann, Cornelius, and Engin Bozdag. 2014. "Staking Out the Unclear Ethical Terrain of Online Social Experiments." *Internet Policy Review* 3 (4). doi:10.14763/2014.4.338.

Reiter, Jerome P. 2012. "Statistical Approaches To Protecting Confidentiality For Microdata And Their Effects On The Quality Of Statistical Inferences." *Public Opinion Quarterly* 76 (1): 163–81. doi:10.1093/poq/nfr058.

Reiter, Jerome P., and Satkartar K. Kinney. 2011. "Sharing Confidential Data for Research Purposes: A Primer." *Epidemiology* 22 (5): 632–35. doi:10.1097/EDE.0b013e318225c44b.

Resnick, Brian. 2016. "Researchers Just Released Profile Data on 70,000 OkCupid Users Without Permission." *Vox*. http://www.vox.com/2016/5/12/11666116/70000-okcupid-users-data-release.

Restivo, Michael, and Arnout van de Rijt. 2012. "Experimental Study of Informal Rewards in Peer Production." *PLoS ONE* 7 (3): e34358. doi:10.1371/journal.pone.0034358.

Riach, P. A, and J. Rich. 2002. "Field Experiments of Discrimination in the Market Place." *The Economic Journal* 112 (483): F480–F518. doi:10.1111/1468-0297.00080.

Riach, Peter A., and Judith Rich. 2004. "Deceptive Field Experiments of Discrimination: Are They Ethical?" *Kyklos* 57 (3): 457–70. doi:10.1111/j.0023-5962.2004.00262.x.

Rich, Judith. 2014. "What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted Since 2000." SSRN Scholarly Paper ID 2517887. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2517887.

Richman, Josh. 2015. "Stanford and Dartmouth Researchers Broke Law with Election Mailer, Montana Official Says." *San Jose Mercury News*, May. http://www.mercurynews.com/nation-world/ci_28100916/stanford-and-dartmouth-researchers-broke-law-election-mailer.

Rijt, Arnout van de, Soong Moon Kang, Michael Restivo, and Akshay Patil. 2014. "Field Experiments of Success-Breeds-Success Dynamics." *Proceedings of the National Academy of Sciences* 111 (19): 6934–9. doi:10.1073/pnas.1316836111.

Robinson, Walter M., and Brandon T. Unruh. 2008. "The Hepatitis Experiments at the Willowbrook State School." In *The Oxford Textbook of Clinical Research Ethics*, edited by E. J. Emanuel, R. A. Crouch, C. Grady, R. K. Lie, F. G. Miller, and D. Wendler, 386–97.

Sandvig, Christian, and Karrie Karahalios. 2016. "Most of What You Do Online Is Illegal. Let's End the Absurdity," June. https://www.theguardian.com/commentisfree/2016/jun/30/cfaa-online-law-illegal-discrimination.

Schauer, Frederick. 1978. "Fear, Risk and the First Amendment: Unraveling the Chilling Effect." *Boston University Law Review* 58: 685. http://heinonline.org/HOL/Page?handle=hein.journals/bulr58&id=695&div=&collection=.

Schechter, Stuart, and Cristian Bravo-Lillo. 2014. "Using Ethical-Response Surveys to Identify Sources of Disapproval and Concern with Facebook's Emotional Contagion Experiment and Other Controversial Studies." *Microsoft Research Technical Report* MSR-TR-2014-97 (October). http://research.microsoft.com/pubs/220718/CURRENT%20DRAFT%20-%20Ethical-Response%20Survey.pdf.

Schneider, Carl E. 2015. *The Censor's Hand: The Misregulation of Human-Subject Research*. 1 edition. Cambridge, Massachusetts: The MIT Press.

Schrag, Zachary M. 2010. *Ethical Imperialism: Institutional Review Boards and the Social Sciences, 1965-2009*. 1 edition. Baltimore: Johns Hopkins University Press.

———. 2011. "The Case Against Ethics Review in the Social Sciences." *Research Ethics* 7 (4): 120–31. doi:10.1177/174701611100700402.

Schultze, Ulrike, and Richard O. Mason. 2012. "Studying Cyborgs: Re-Examining Internet Studies as Human Subjects Research." *Journal of Information Technology* 27 (4): 301–12. doi:10.1057/jit.2012.30.

Schwartz, Paul M., and Daniel J. Solove. 2011. "The PII Problem: Privacy and a New Concept of Personally Identifiable Information." SSRN Scholarly Paper ID 1909366. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=1909366.

Selinger, Evan, and Woodrow Hartzog. 2015. "Facebooks Emotional Contagion Study and the Ethical Problem of Co-Opted Identity in Mediated Environments Where Users Lack Control." *Research Ethics*, May, 1747016115579531. doi:10.1177/1747016115579531.

Seltzer, William, and Margo Anderson. 2008. "Using Population Data Systems to Target Vulnerable Population Subgroups and Individuals: Issues and Incidents." In *Statistical Methods for Human Rights*, edited by Jana Asher, David Banks, and Fritz J. Scheuren, 273–328. Springer New York. http://link.springer.com/chapter/10.1007/978-0-387-72837-7_13.

Shaw, David. 2015. "Facebooks Flawed Emotion Experiment: Antisocial Research on Social Network Users." *Research Ethics*, May, 1747016115579535. doi:10.1177/1747016115579535.

Sheehan, Mark. 2011. "Can Broad Consent Be Informed Consent?" *Public Health Ethics* 4 (3): 226–35. doi:10.1093/phe/phr020.

Singal, Amit G., Peter D. R. Higgins, and Akbar K. Waljee. 2014. "A Primer on Effectiveness and

Efficacy Trials." *Clinical and Translational Gastroenterology* 5 (1): e45. doi:10.1038/ctg.2013.13.

Singel, Ryan. 2009. "Netflix Spilled Your Brokeback Mountain Secret, Lawsuit Claims." *Wired.* http://www.wired.com/2009/12/netflix-privacy-lawsuit/.

Sinnott-Armstrong, Walter. 2014. "Consequentialism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2014. http://plato.stanford.edu/archives/spr2014/entries/consequentialism/.

Soeller, Gary, Karrie Karahalios, Christian Sandvig, and Christo Wilson. 2016. "MapWatch: Detecting and Monitoring International Border Personalization on Online Maps." In *Proceedings of the 25th International Conference on World Wide Web*, 867–78. WWW '16. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee. doi:10.1145/2872427.2883016.

Solove, Daniel J. 2010. *Understanding Privacy.* 2/28/10 edition. Cambridge, Mass.: Harvard University Press.

Sommer, Robert. 1969. *Personal Space: The Behavioral Basis of Design.* Englewood Cliffs, N.J.: Prentice Hall Trade.

Sommers, Roseanna, and Franklin G. Miller. 2013. "Forgoing Debriefing in Deceptive Research: Is It Ever Ethical?" *Ethics & Behavior* 23 (2): 98–116. doi:10.1080/10508422.2012.732505.

Spitz, Vivien. 2005. *Doctors from Hell: The Horrific Account of Nazi Experiments on Humans.* Sentient Publications.

Stark, Laura. 2012. *Behind Closed Doors: IRBs and the Making of Ethical Research.* 1 edition. Chicago ; London: University Of Chicago Press.

Sunstein, Cass R. 2002. "Probability Neglect: Emotions, Worst Cases, and Law." *Yale Law Journal* 112 (1): 61–107. http://www.yalelawjournal.org/essay/probability-neglect-emotions-worst-cases-and-law.

———. 2005. *Laws of Fear: Beyond the Precautionary Principle.* Cambridge, UK ; New York: Cambridge University Press.

Sweeney, L. 2002. "K-Anonymity: A Model for Protecting Privacy." *International Journal on Uncertainty Fuzziness and Knowledgebased Systems* 10 (5): 557–70.

Sweeney, Latanya, Mercè Crosas, and and Michael Bar-Sinai. 2015. "Sharing Sensitive Data with Confidence: The Datatags System." *Technology Science*, October. http://techscience.org/a/2015101601/.

Tene, Omer, and Jules Polonetsky. 2016. "Beyond IRBs: Ethical Guidelines for Data Research." *Washington and Lee Law Review Online* 72 (3): 458. http://scholarlycommons.law.wlu.edu/wlulr-online/vol72/iss3/7.

Tockar, Anthony. 2014. "Riding with the Stars: Passenger Privacy in the NYC Taxicab Dataset." *Neustar Research.* http://research.neustar.biz/2014/09/15/riding-with-the-stars-passenger-privacy-in-the-nyc-taxicab-

Vaccaro, K., K. Karahalios, C. Sandvig, K. Hamilton, and C. Langbort. 2015. "Agree or Cancel? Research and Terms of Service Compliance." In *ACM CSCW Ethics Workshop: Ethics for Studying Sociotechnical Systems in a Big Data World.*

Vaughan, Ted R. 1967. "Governmental Intervention in Social Research: Political and Ethical Dimensions in the Wichita Jury Recordings." In *Ethics, Politics, and Social Research*, edited by Gideon Sjoberg, 50–77. Schenkman Publishing Comany.

Verma, Inder M. 2014. "Editorial Expression of Concern: Experimental Evidence of Massivescale

Emotional Contagion Through Social Networks." *Proceedings of the National Academy of Sciences* 111 (29): 10779. doi:10.1073/pnas.1412469111.

Warren, Samuel D., and Louis D. Brandeis. 1890. "The Right to Privacy." *Harvard Law Review* 4 (5): 193–220. doi:10.2307/1321160.

Wendler, David, Leah Belsky, Kimberly M. Thompson, and Ezekiel J. Emanuel. 2005. "Quantifying the Federal Minimal Risk Standard: Implications for Pediatric Research Without a Prospect of Direct Benefit." *Journal of the American Medical Association* 294 (7): 826–32. doi:10.1001/jama.294.7.826.

Wesolowski, Amy, Caroline O. Buckee, Linus Bengtsson, Erik Wetter, Xin Lu, and Andrew J. Tatem. 2014. "Commentary: Containing the Ebola Outbreak - the Potential and Challenge of Mobile Network Data." *PLoS Currents*. doi:10.1371/currents.outbreaks.0177e7fcf52217b8b634376e2f3efc5e.

Wimmer, Andreas, and Kevin Lewis. 2010. "Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook." *American Journal of Sociology* 116 (2): 583–642. http://www.jstor.org/stable/10.1086/653658.

Windt, Peter Van der, and Macartan Humphreys. 2016. "Crowdseeding in Eastern Congo Using Cell Phones to Collect Conflict Events Data in Real Time." *Journal of Conflict Resolution* 60 (4): 748–81. doi:10.1177/0022002714553104.

Wu, Felix T. 2013. "Defining Privacy and Utility in Data Sets." *U. Colo. L. Rev.* 84: 1117. http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/ucollr84&section=40.

Yakowitz, Jane. 2011. "Tragedy of the Data Commons." *Harvard Journal of Law & Technology* 25: 1. http://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/hjlt25&section=4.

Yoon, Carol Kaesuk. 1997. "Families Emerge as Silent Victims Of Tuskegee Syphilis Experiment." *The New York Times*, May. http://www.nytimes.com/1997/05/12/us/families-emerge-as-silent-victims-of-tuskegee-syphilis-experiment.html.

Zevenbergen, Bendert, Brent Mittelstadt, Carissa Véliz, Christian Detweiler, Corinne Cath, Julian Savulescu, and Meredith Whittaker. 2015. "Philosophy Meets Internet Engineering: Ethics in Networked Systems Research. (GTC Workshop Outcomes Paper)." SSRN Scholarly Paper ID 2666934. Rochester, NY: Social Science Research Network. http://papers.ssrn.com/abstract=2666934.

Zimmer, Michael. 2010. "But the Data Is Already Public: On the Ethics of Research in Facebook." *Ethics and Information Technology* 12 (4): 313–25. doi:10.1007/s10676-010-9227-5.

———. 2016. "OkCupid Study Reveals the Perils of Big-Data Science." *WIRED*. https://www.wired.com/2016/05/okcupid-study-reveals-perils-big-data-science/.

Zimmerman, Birgitte. 2016. "Information and Power: Ethical Considerations of Political Information Experiments." In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, edited by Scott Desposato. Routledge.