

Afterword: Ethics as Impact—Moving From Error-Avoidance and Concept-Driven Models to a Future-Oriented Approach

Social Media + Society
July–September 2018: 1–11
© The Author(s) 2018
DOI: 10.1177/2056305118784504
journals.sagepub.com/home/sms


Annette N. Markham 

Abstract

Beginning with the premise that most regulatory guidelines for research ethics are deeply flawed, this article walks the reader through three models that can help social researchers, technologists, and designers identify and reflect on how they're approaching ethics, or "doing the right thing" in their own work. The first, an error-avoidance model, has traditionally focused on creating frameworks to help researchers avoid repeating historical ethical violations. The second concept-driven model focuses on refining the concepts that undergird core ethical frameworks. Both are dominant in human-focused research and tend to be highly proceduralized, implemented a priori or from the top down as part of largescale regulatory structures. In both of these models, the agency of the researcher is removed or dismissed as less relevant than the agency of the system. The article draws on recent controversies around data collection and corporate experimentation on social media users as well as two academic research cases to illustrate how these two models fail repeatedly because they do not retain enough flexibility to allow for recontextualizing ethics as needed on a case-by-case basis. The third, an impact-model of ethics, offers an alternative whereby researchers, technologists, and designers can take a more active role in decisions about the contexts they study, by exploring the possible positive and negative impact of their work. This article invites us to work toward building a different balance in agential distribution in our models around responsible conduct of research, so that the conceptual and regulatory systems that guide and impede our actions are more balanced with our own agency as decision makers, with accountability and responsibility for doing the 'right thing'.

Keywords

ethics, method, social media, impact, contextual ethics

Introduction

This short primer on ethics and digital research emerges from the contention that institutional regulatory models that guide research ethics are deeply flawed and broken. It aligns with a now longstanding body of scholarship that critically analyzes and resists the decontextualization of ethics in philosophy or the human sciences and also the bureaucratized treatment of ethics by regulatory bodies. Using specific cases, this article illustrates how error-avoidance and concept-driven models, which lie at the heart of all regulatory guidelines and procedures, fail repeatedly in contemporary research situations. The ideas in this article are shaped in academic environments for social research, but clearly extend beyond these walls, as the problems around the consideration of ethics are being experienced across industries working with or handling personal data.

Before proceeding, it should be noted that around the world, awareness of this problem is growing; research funding

agencies, governmental oversight bodies, and institution-specific ethics review boards are adapting to the changes brought about by digital media and the Internet. There is a growing recognition that the relationship between ethics review boards and researchers ought to be one of continuing, cooperative dialogue. Still, the general pattern for review and approval of research remains strongly top down, bureaucratic, and locked in conceptualizations of research that do not fit well with contemporary complexities of research environments. Perhaps because systems are slow to change, a feeling of antagonism

Aarhus University, Denmark

Corresponding Author:

Annette N. Markham, School of Communication and Culture, Aarhus University, Helsingørsgade 14, 8200 Aarhus N, Denmark.
Email: amarkham@gmail.com



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

persists. Especially in countries with heavy top down regulatory systems, such as the United States, Canada, or Australia, this feeling reinforces the practice of writing research proposals in ways that allow them to be exempted from review or expedited through approval processes. This can be counterproductive if one's research really needs guidance. Worse, it leads to situations whereby researchers conceptualize ethics as a list of checkboxes on a form, or limit their ethical consideration to only those items that have been previously identified as crucial and therefore appear in regulatory documents. The systemic problem in heavily regulated environments is that for many researchers "ethics" is a set of procedural considerations taken care of prior to the conduct of a study, rather than a supple practice of reflexivity about the emergent and ongoing impact of research practices as they unfold in situ. Yet, ethical dilemmas have a way of creeping up on us, often becoming apparent only after something goes wrong. Therefore, even as conceptual and regulatory development around ethics and digital media research has been strong¹, researchers—and regulators—continue to grapple with ambiguous and contested notions of data protection and management, intellectual property, privacy and personal information, vulnerability of participants, consent, responsibility, accountability, not to mention individual or social harms.

This primer categorizes ethical frameworks in three ways to highlight different starting points for ethical considerations by researchers. These frameworks are as follows: the *error-avoidance* model, the *concept-driven* model, and the *impact-focused* model. The first two models characterize the dominant conceptualization of ethics for research practice. The third, an impact model for ethics, is a future-oriented perspective that counters some of the weaknesses of the first two. I will first explore the problems inherent in the error-avoidance and concept-driven modalities in relation to two specific examples of seemingly ethical research practices gone awry. I will then attempt to describe and elaborate the possibilities for an impact-focused model of research ethics.

Shifting one's lens can help bring the important principles back to central focus. This, I argue, requires moving away from the effort to more precisely define these concepts to instead consider a different set of concepts and practices more related to the *processes through which research is conducted*. 'Process' here refers to three different contextual levels that get collapsed when we think about research processes: the historical context, the institutional context, and the research situation. The third model, the *impact-focused* model of ethics (Markham, 2015), gives more agency to the researcher, who makes decisions in specific research situations. At the level of the research situation, we can pay attention more directly and appropriately to action and reaction, responsibility and accountability. In other words, instead of fitting actions to concepts or regulations, we can focus on the actual and possible consequences of our actions. This impact approach is decidedly future-oriented, whereas error-avoidance and concept-driven models privilege past practice and historical precedent as the dominant arbiters of the ethical.

To illustrate the importance of this perspectival shift in temporal focus from error and/or concept to impact, I use two examples, both provided by Swedish technofeminist and ethics scholar Eva Svedmark (2016). These illustrate how responsible conduct of research occurs beyond the limited scope of regulatory guidelines, which is particularly necessary for research involving networked media and the use of digital social data, as these situations present still-novel and unexpected challenges and potential harms. Digital media forms and practices entail distinctive social ontologies that require an impact-driven ethics.

These cases also emphasize that ethical dilemmas are just that: Dilemmas. They are often unresolvable, which is a difficult challenge for regulatory agencies and research communities to accept. For example, one of the central ideas behind institutionalized procedures is the avoidance of harm. Yet, we know that harm is almost impossible to predict. Sometimes, harm is inevitable, or has already been done before the researcher arrived. Sometimes harm that could not be anticipated occurs long after the original research is conducted. Sometimes harm cannot be avoided. Sometimes, as many of us have learned through trial and error, the researcher will be harmed in the natural course of the study. How to handle these situations from a moral standpoint often differs from how one might address this from an ethical, regulatory, or legal standpoint. From a more practical method or practice standpoint, most ethically complicated situations pose dilemmas whereby one must make a choice among options with no clear outcome. No matter how much we turn to regulators or guidelines or procedures or others for help, these situations are in the end a judgment call, which is the hallmark of contextual, casuistic, or case-based, approaches.

The Problem With Error-Avoidance Ethical Models

Ethical breaches in history have created a strong reliance on regulation to guide how we think about and treat people as we go about our work in design labs, ethnographic fields, tech companies, or other places where we engage in interventions, experiments, or activities that might cause some sort of harm. Using historical cases to identify possible actions and problems creates what I call "error-avoidance" ethics models.

Misconduct is the most common starting point for discussions of research integrity at the level of academic institutions. These are well intended and have been, for the most part, effective means of creating widely accepted practices and procedures for responsible conduct of research. A key goal is to alert researchers of areas of misconduct so they can avoid making mistakes. Institutions tend to initially focus on mistakes like fabrication of data, plagiarism, and falsification. When applied to research on "human subjects," the guidelines become much more specific.² Before the study

begins, if we have classified the object of analysis as a person, or human subject, we define what constitutes vulnerability and harm. These definitions take a mostly procedural shape as we operationalize them in our research designs, whereby harm is avoided or minimized by gaining informed consent or developing strategies to protect privacy. Therefore, in the social sciences, well intended and defensible ethical stances are generally determined in advance of a study, and applied as a deductive strategy to protect the research, the institution, and the potential participants.

Our initial definitions for the phenomena will significantly shape whether and how we see the “human” in the situation. Defining Twitter as an information stream or a blog as a published text invokes a different set of principles than if they are described as an extension of the person, or a diary posted in a place of perceived privacy. Take the 2014 case of what has become known as the *Facebook Emotional Contagion study*. Facebook deliberately changed the newsfeed for certain users, tweaking the emotionality of posts for a certain period of time and collected anonymized data on usage characteristics during this time, for later analysis. Defining the study as “A/B testing” versus “experiment on users” would yield considerably different ethical parameters.³ In the former, the creators of a service, product, or message (like an advertisement) might test versions A and B on customers, clients, or audiences, to determine which version more effectively yields the desired result. Using this logic, one could say Facebook developers were testing content on users, giving them different news to ascertain how it might make them feel, in order to improve the service in the future. This is, however, not how many people felt when learning about the study. The general outrage following the publication of findings from this study⁴ leads us to believe that by and large, people believed they might have been the unwitting and nonconsenting subjects of experimentation, exposed to different stimuli to test their responses.

A good researcher will meet rising ethical dilemmas as they emerge, by assessing the working assumptions and making necessary shifts in conduct based on situationally based judgment or wisdom that emerges from practical experience rather than deduced from a priori concepts or categories (phronesis).⁵ But if there is no obvious disjuncture or dilemma, ethical guidelines can and often are applied in an unproblematic way. In many cases, the outcome of one’s research does not harm the community, the participant, or the researcher. Because harm cannot often be ascertained in advance of a situation, problems can arise in unexpected ways. Let’s look at one such case.

Case 1: When Doing the Right Thing (Informed Consent) Causes the Wrong Thing

To do the right thing does not always involve doing what common ethics models or regulators consider the right thing.

The following is a good example of how following an error-driven ethics model caused more harm than good.

The researcher⁶ had been following a “Pro Ana” blog written by a young Swedish woman to talk about her eating disorder. The blog was written under a pseudonym and the woman was careful never to say anything that could reveal her offline identity. The researcher had been thinking about whether this blog might be a good fit for a larger research project on emotional online sharing, so while she was paying attention to this blog as a casual follower, she was not doing any official research about it yet.

Because the situation involved ethically sensitive information, the researcher consulted Swedish regulations for ethical research conduct as well as international, Internet-specific ethical guidelines. Everything the researcher read pointed at getting informed consent in order to carry on the study of this Pro Ana blog. Since the young woman was writing under pseudonym, the researcher contacted the potential participant through her anonymous email address, explaining the objectives of the study and asking for informed consent to study her blog. The woman replied swiftly, indicated she understood, gave informed consent, and indicated she was delighted to be a part of this project.

Very soon after this email exchange, the participant’s blog posts changed dramatically. She appeared more intensely interested in her self starvation. Her posts included statements such as “*Finally I have the motivation to get down to a BMI of 14*” and “*It’s a great motivation to not eat now that I have someone watching over me.*” This type of posting became more intense and the eating disorder had a more outspoken voice than ever. The blogger also stated that she felt like she was finally ready to stop eating altogether.

The researcher interpreted this as evidence that the blogger was suddenly very aware of and influenced by the researcher’s presence and worse, that the blogger had become increasingly obsessed with her eating disorder. It seemed that the researcher’s presence triggered her illness in a very real way. *Informed consent seemed to be the starting point for this negative change.*

The researcher felt compelled to try to correct this situation and contacted the participant again. She underscored how she was not interested in the participant’s eating disorder but in the expression of emotions and feelings through blogging. Soon after this message was sent, the participant changed her email address and blocked access to the blog with a password. Essentially, the researcher was thrown out. Without any further information, the researcher could only assume her actions had caused both situations. She was greatly shaken by this experience.

This case provides important evidence that often good ethics only *incidentally* align with legal parameters and regulatory norms in the first place. The researcher had done everything right, according to both national and international recommendations and regulations. But from all apparent outcomes, seeking informed consent actually caused harm.

Informed consent is conceptualized in different ways globally, but almost always, there is a core assumption that if informed consent is obtained, potential harm will diminish. This is a faulty assumption because it requires some connection between consent and harm in the first place, which is rarely if ever the case.

This requires some picking apart: The motive behind informed consent is that a subject or participant should be allowed and enabled to understand fully what the research is about and what their role as participant implies. It is implied that gaining their consent will preserve the the participant's autonomy to make decisions about whether to consent or not. Asking for consent will also show respect to the person. Importantly, the argument is built from previous error-cases, where subjects were not asked whether they consented or their consent was not informed. In Stanley Milgram's obedience studies, for example, participants consented but were misled about what the studies entailed. In the Tuskegee Syphilis studies, some participants lost the opportunity to receive treatments, with life-threatening consequences. In both cases, informed consent would have allowed persons more autonomy and choice and would have enabled them to opt out of a study or adequately understand the risks involved. But informed consent would not have reduced harm, necessarily.

The described case with the pro-ana blogger, like many others, demonstrates that the consequences of one's research activities, in terms of felt/actual harm, can be opposite than anticipated. Part of the problem is certainly bound up in the way regulatory guidance is formulated. By focusing on the avoidance of errors that have occurred previously, one's attention is drawn to how those particular errors might occur, at the expense of other possible problems. It is by now a common argument among ethics scholars that flexible adaptation is necessary and that a primary failure of the system is its continued top-down approach. Here, we can see that an error-avoidance approach can hamper one's sensitivity to the context and ability to make or alter ethical decisions when needed.

Acknowledging that there are significant problems with the historical error-based notion of informed consent in digital social contexts, there have been continuing efforts to expand what it means and how it is obtained. Tiered consent, revisiting consent, waiving consent when impossible to obtain, and finding different ways of documenting consent have become useful contemporary techniques. However, none of these resist the underlying logic that repeatedly assures us that "informed consent" is *always* the *best* concept through which researchers can protect autonomy. Herein lies another flaw; this logic compels researchers to focus on a narrow set of assumptions about what counts as respect and autonomy and based on that, an equally narrow set of actions to enable or ensure these principles in the course of a study. Below, I continue this point with a second case.

The Problem With Concept-Driven Ethics

Our concepts of subjectivity, presence, place, and privacy—and how all of the phenomena cohere in the figure of the "person" (the true and only legitimate source of consent in dominant ethical protocols)—have been confounded by the technologies surrounding the Internet. A strong line of scholarship emerged in the late 1990s that critically examined and sought to refine the core concepts around research ethics. As used here, concepts refer to the operationalization of core human rights principles. What we have found in the past 20 years of conceptual work is that however much one might want to find concrete solutions and tools to guide ethical decision making, the process of standardizing concepts vis-à-vis regulatory statements results in presumptive, regulation-driven classification of specific contexts, which limits nuanced and contingency-based decision making.

Take the concept of privacy: In trying to fix the obviously broken regulatory guidelines for determining what is private and how one's privacy should be protected, we have witnessed nearly two decades of efforts to articulate what this concept means in the digital age (see the 2002 and 2012 AoIR ethics guidelines for summaries). The effort to determine how we might assess and therefore protect what is considered private is a laudable goal. But the concept cannot contain the multitude of definitions-in-context that are used by everyday people to make sense of their relationship with their own data. Social media researchers and designers struggling with the practical and methodological challenges surrounding the issue of privacy find themselves on constantly shifting conceptual terrain. Yet in popular media and most regulatory documents, "privacy" continues to be addressed as a binary that contrasts private with public.

Individual and cultural definitions and expectations of privacy are ambiguous, contested, and changing. Commonly, people operate in public online spaces but maintain strong expectations of privacy. Swedish scholars Svedmark and Nyberg (2009) interviewed dozens of Swedish youth (aged 15–25 years) who posted revealing pictures of themselves in public online spaces (ref). Their responses revealed a complex and, perhaps for some readers, a counterintuitive understanding of privacy. As one participant put it, "I have this blog online because it's the only place where i feel safe and be sure my parents won't find it. If it had it in a diary in my room, they would find it."

Perceived versus actual privacy is a significant conceptual development that has guided much of our current thinking about privacy and ethical social Internet research (King, 1996)⁷ e.g., Bromseth, 2003. While this distinction is a useful progression from the binary of public and private, it remains problematic in application because it pre-defines and makes a priori that which is actually emergent and after the fact. It is quite common that people acknowledge that the substance of their communication is public but feel dismayed

when the information is taken out of context or used in a way that is unexpected or unwanted. Youth interviewed by Alice Marwick and danah boyd (2014) felt angry that their Facebook information was used by teachers and administrators in a school-wide public presentation to illustrate the dangers of posting private information in public spaces. They knew the information was relatively public. It was the manner in which it was used that made them feel violated (p. 6). And their assessment of privacy at that point was after the fact, after it had been used out of context.

Privacy may not be the most relevant concept at all in a digital age, but rather control and context. As Nissenbaum (2009) has pointed out, “what people care most about is not simply restricting the flow of information but ensuring that it flows appropriately” (Intro, para. 2). So to consider how we might protect participants, we researchers need to also look toward a different set of issues surrounding the situatedness of the self. When we consider how people talk about their experiences in situations where privacy has been breached, a more apt question might be, “What is a person’s relationship to their information?” or more specifically focused on relationality, even: “How does a person feel about his or her relationship to his or her information?” (Markham, 2012). In *It’s Complicated*, danah boyd’s (2014a) participants emphasize in multiple ways how privacy is not just complicated, but idiosyncratic, individual, contextual, and relational.

The problem with concepts (or more precisely, the logic of conceptualization) is that in the process of becoming strong, they become reified, incorrigible, and resistant to modification based upon present experience and future possibilities. Concepts thus may seem to preexist as templates to help us enter and describe situations, but in actuality, they are continually emerging from actual situations, at least if we are talking about concepts that prompt ethical consideration from scientists. When we apply privacy as a predefined concept, our definition might align with what is salient in the moment, but we also might end up attending to the wrong ethical principle—wrong in this sense meaning less salient or important to the situation. Below, I offer a second case that demonstrates how this can happen quite easily, especially if the research situation seems straightforward.

Case 2: When Regulatory Definitions of Vulnerability Are Exactly Wrong

A researcher studying bloggers’ descriptions of their everyday life needed to decide whether or not to quote directly from the blogs or disguise them through the use of pseudonyms. All the blogs were public, but the researcher knew from previous studies on perceived privacy that just because the blogs were publicly accessible did not mean they were perceived by their creators as public. While she would contact the bloggers and ask them whether they wanted to remain anonymous, part of the researcher’s thinking was dictated by Swedish ethics regulations’ conceptualization of

vulnerability. As in most countries, the ethical regulations in Sweden specify the concept of “vulnerability.” Swedish law clearly states that extra care and privacy protection must be taken in study of children, or when the research covers topics related to mental or physical illness, religion, ethnicity, or sexuality.

The category of vulnerability, as operationalized by most international regulatory guidelines, certainly applied to one blogger in the study: Kersten. Kersten blogged about the loss of her daughter who had committed suicide. The posts on Kersten’s blog were personal and revealing, full of details about her emotional state of mind. The grief was so present in the blog that you could almost touch it. For an entire year after her daughter died, Kersten took readers on a daily retrospective journey back in time, sharing detailed descriptions of each day of the full year leading up to the day her daughter decided to end her life. Kersten did not blog anonymously and shared her own as well as the deceased daughter’s identity. Following Swedish guidelines, the researcher labeled Kersten a vulnerable participant. Her deep grief, frequent expressions of depression, and topic of blogging were all taken to be clear indicators that she requires extra care as the topic of and/or participant in social research.

In contrast, a second blogger, Greta, was part of a community of women interested in home styling and interior design. She wrote and posted pictures about her everyday life in her blog and shared the blog with the larger community of women. The style of the blog was open, friendly, and happy. Greta wrote under a pseudonym like many of the other women in this community, but sprinkled her blog posts liberally with photos of her home and garden. The researcher did not label Greta a vulnerable participant.

Looking at these blogs from the outside, it is easy to place Kersten and Greta in different predetermined categories for vulnerability. But as it turned out, the researcher’s categorization was misplaced in both cases. When asked, Kersten was unconcerned by the idea of being studied. She described herself as being in the middle of a life crisis with high emotional turmoil. But at the same time, she insisted she wanted to make her voice heard. Kersten wanted to put suicide in the spotlight through her blog and thereby raise awareness about how suicide affects loved ones left behind. She requested that she be named and that her publicity be protected, not her privacy.

Greta, on the contrary, expressed a great deal of uncertainty and fear about being the subject of study. She was concerned that if published as part of a study, she would lose her anonymity. She also expressed a feeling that her privacy was being invaded. It turns out Greta’s blog was not known to her friends or family nor to any one at work. While the content might not be sensitive, Greta wanted her blog to remain private and she did not want it to be known by people in her physical surroundings. In the researcher’s words, Greta was adamant that she did not want to become “the blogger” at work and she did not want to share her private life and home with colleagues and distant acquaintances.

This case illustrates how focusing on the operationalized concept can potentially mislead researchers into thinking the situation is clear when in reality, it is not. Greater clarity may emerge from focusing on the uniqueness of the context under study rather than the supposed similarities of the present case to prior cases or categorizations.

As a side note, one might make a strong argument that Kersten is actually vulnerable even if she insists she is not. Perhaps, Kersten has not fully considered the implications of being studied and publicly cited because she is in a highly emotional state. Also, there is a strong case to suggest that even if Greta does not fall into the regulatory category of vulnerability, her privacy should be protected by default because she blogged under a pseudonym, which indicates her desire to remain unknown or anonymous.

Regardless of alternate explanations or different choices about who is actually vulnerable, the point is that it is only through more investigation of the specific research situation and, if relevant, persons under study that we begin to learn what vulnerability means, to this or that person, or in this or that moment. This case illustrates how *reliance on concepts rather than contextual details can lead not just to misdefining the vulnerability of participants, but to mistaking predetermined notions with actual experiences of harm*. It helps us see how concept-driven ethical stances focus on the concept, in this case vulnerability, whereas an impact-oriented or future-oriented ethical stance would focus the more basic issue of actual or potential harm, which could yield a more appropriate line of questioning for participants and lead the researchers to understand better what might be required to protect them.

People have extremely unpredictable notions about privacy, vulnerability, anonymity, and harm. In all research, rigorous examination of the ethical challenges before starting the study is a good idea. But in the examples above, relying on the concepts—as defined and regulated by governing bodies—proved to be a mistake. It illustrates how the conceptual frameworks are as muddled as the regulatory ones, as James Moor noted in 1985 and Michael Zimmer reminds us of in 2016. Rather than trying to find the best conceptual framework that can guide regulatory models, it might be more sensible to consider that ethical concepts have always been muddled and therefore “doing the right thing” requires continual adjustment in situ as they are operationalized into decisions that have consequences, small and large.

Moving Toward an Impact Model of Ethics

The two cases above demonstrate the challenge of determining the best actions in a specific situation when one’s decisions are limited by disciplinary norms or regulatory directives. Indeed, in arenas where heavy regulatory mechanisms are present (i.e., US or Australian academic institutions), a researcher’s agency as well as responsibility is narrowly assigned to the proper enactment of the regulation rather than achievement of the principle. This occurs through several layers of translation: When

the principle of “autonomy” gets operationalized as “informed consent,” it gets further operationalized as the regulations are spelled out—Informed consent is the outcome of a specified set of procedures to inform a participant who is considering participating in a study about the specific risks and benefits and to ensure the participant fully comprehends the information they have been given and further understands what consent entails. Proceduralizing these actions is a late stage translation whereby the accomplishment of informed consent is specified and routinized, such as obtaining a signature from the participant. This is, incidentally, also where we can see small regulatory adjustments to make the procedures fit new scenarios (e.g., from physical signature to electronic consent), which can be quite useful, but can obscure deeper problems in the baseline logics that make the adjustment necessary in the first place.

Rote adherence to regulatory requirements is a typical behavior in heavily regulated institutions, exacerbated by the fear that if one does not do exactly as required, the ethics committee will not approve the research and your study will stall. It is no wonder researchers operating under such institutionalized strictures do not go beyond the bounds of what is required, either to resist the procedures when they do not fit the context or to go above and beyond the requirements when the situation calls for extra ethics.

The flipside of the innocent or perhaps somewhat mindless adherence to rules is the willful use of the regulatory framework to justify making a bad decision. In 2016, for example, Emil Kirkegaard harvested personal data from over 70,000 users of the dating site OKCupid. After reformatting the data, he uploaded it online to a data sharing site for other researchers to download and use. The information in the data set included personal details like age, gender, location, username, sexual preference, as well as user’s answers to thousands of questions asked by OKCupid. The data release prompted a storm of responses from the international community—scholars, pundits, politicians, other data scientists, philosophers—about informed consent, privacy, terms of service, educational responsibilities, and other ethics-related concerns. Of particular interest here is how a procedural justification, or “letter of the law” argument, was used repeatedly to attempt to justify a poor⁸ ethical decision.

As Kirkegaard wrote before the data were released, in a paper that has now been removed, 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.

After Kirkegaard announced the release of the data on Twitter, he responded to various criticisms in a similarly straightforward way. In one of the replies, someone asked, “This data set is highly re-identifiable. Even includes usernames? Was any work at all done to anonymize it?” Kirkegaard responded, “No. Data is already public.” To the question, “Interesting that the data was scraped and not provided by OKCupid. Are they ok with that?” Kirkegaard replied, “Don’t know, don’t ask :).” As outrage around the

case grew, the data were eventually removed from the Open Science Foundation.⁹ OKCupid claimed the act of scraping violated their terms of service. But the damage had been done because the data had already been spread.

These data breach helped cement growing concerns about how existing regulations can unintentionally function (or be invoked deliberately) in ways that are completely contrary to the human rights principle they are intended to protect. Even after a widespread international critical response, Kirkegaard continued to reference procedures and rules rather than the concepts or principles underlying the construction of these legal parameters. He later told *Fortune*,¹⁰ “We thought this was an obvious case of public data scraping so that it would not be a legal problem.”

The OKCupid data release, the recent Cambridge Analytica case,¹¹ and many other data-related mishaps, deliberate or otherwise, have prompted important efforts to connect data more closely with persons. In fact, we can see the vocabulary within a larger community of practice changing: In 2012, the Association of Internet Researchers’ working committee on ethics argued ethics guidelines that data always involve persons, saying, “because all digital information at some point involves individual persons, consideration of principles related to research on human subjects may be necessary even if it is not immediately apparent how and where persons are involved in the research data” (Markham & Buchanan, 2012, p. 4). Metcalf and Crawford (2016) bring this relationship even closer when they suggest replacing human subject with “data subject” or “data subjectivity.” Zook et al. (2017) come straight to the point, saying the first rule of responsible big data research is to “acknowledge that data are people and can do harm” (n.p.).

It is clear from these and other discussions that many individuals across broad communities of practice around data science, digital culture, and tech design understand that the entanglement of humans and their data is complicating how one should conduct research. However, it can be difficult to keep this entanglement in mind when data are removed from its source, which can happen in many ways, such as through data aggregation, data cleaning/stripping, passing data through several sources, or even the passage of time. For example, blood samples extracted from the Havasupai tribe

in 1989 lay frozen in an obscure corner of a lab at Arizona State University for more than a decade before they were re-analyzed with a new set of hypotheses. Innocent-seeming research questions, applied to the data, yielded findings that caused an entire community to feel psychological and cultural harm (Harmon, 2010; Sterling, 2011). Elizabeth Buchanan and I refer to a *distance principle* to describe how experiential distance between data and the humans that produce it can obscure potential ethical issues.¹²

An impact model of ethics shifts attention to possible future trajectories. While understanding that it is always useful for researchers to distinguish and clarify the object of study (e.g., data, persons), the characteristics of data (e.g., public/private, raw/cleaned, sensitive/nonsensitive), and the characteristics of people (e.g., who might want their privacy or publicity protected, who are vulnerable, who will use our designs, who are behind the data), an impact model highlights this decision making as a process occurring within larger systems of action and reaction, across longer timespans. With a speculative lens situated at a future point in time, one can look backward, to consider how this particular effect might have been caused. As possible chains of events are laid out, they can be explored in turn, as possible, probable, unlikely, ridiculous, creepy, acceptable, and so forth.

Regardless of what might actually occur or whether one’s analysis is accurate, this analytical process has an anticipatory function. This is particularly useful in large scale data analytics, experimental research, or technology design environments when the ethical problem is not apparent in the immediate moment but can be seen through a speculative lens. This sort of reverse engineering “What happened to get us here?” from a future orientation lays widens the scope of possible outcomes, which can help identify existing blind spots and consider possible side effects, downstream impact, or broader ecological outcomes.

In 2015, Janice Tsai, Sumit Basu, two Microsoft researchers, and I brainstormed what an impact model of ethics might look like. It has strong roots in the prospective ethics of traditional computing and engineering sciences, where one considers the potential outcomes. Some possible scales of impact are offered in Figure 1.

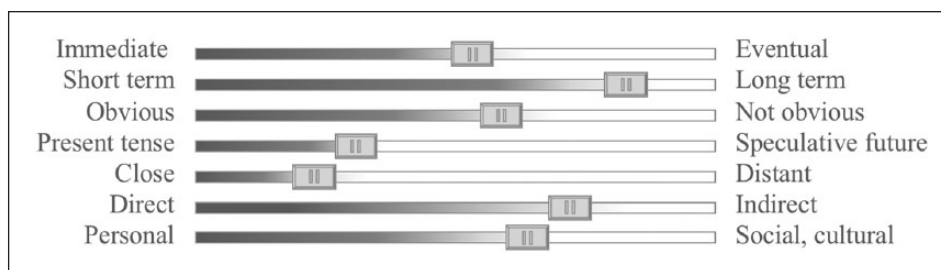


Figure 1. Depicts some possible scales to consider ethical impact of research design. Levels indicate this assessment is case-by-case and adjustable.

In addition to considering impact along different scales, we can separate impact into different arenas. Considered separately, these arenas can help avoid conflating the many factors and agencies and levels of impact within larger ecologies; impact might occur in digital media and data-related research. As with most categorizing schema, the four arenas below are neither mutually exclusive nor all inclusive, but generative; intended as a tool to think with:

Impact Arena 1. Treatment of People

This arena builds on the strengths of psychological and sociological domains, where focusing on human rights is central. The assumption in this category is that researchers and developers interact with people all the time. While it may be clear in some situations that one is working with a person or “human subject,” in other situations, people can be obscured by labels such as profile, interface tester, use case, end user, data point, ping, or digital signal.

In this arena, we might ask: What is the potential impact of manipulating the user’s environment to test the impact of certain system inputs; e.g., changing “facts” in order to test learning software on learners in an actual online course or MOOC; building system pings to prompt Mechanical Turkers to work faster? What is the potential harm of psychological interventions, such as experimenting with online triggers without debriefing, or using deception and manipulation in social research encounters? What is the impact of seemingly innocuous design? For example, what is a possible impact of involving an anonymous individual in a highly sensitive interview or experiment online but they depart midway? The setup of allowing such an abrupt opt-out might satisfy informed consent but cause another concern when the participant is deprived of needed debriefing. What happens when the technical option of “share findings” in genetic testing software leads to unpleasant discoveries about family behavior by third parties? What might be triggered in photo elicitation? Or how might piece workers be impacted by feature-finding in machine learning? What is the potential impact of research on one’s own wellbeing, or on broader communities; e.g., when one’s research design poisons the well for future researchers?

Impact Arena 2. Side effects

This arena builds on the strengths of science and engineering domains, where a precautionary principle aids in the assessment of short and long term social and environmental effects of scientific developments. This impact arena assumes that unintended consequences are natural and unavoidable in any research and technology situation. These might become obvious only later. Side effects may be directly linked to one’s research and tech design or a result of multiple factors across platforms. They may also operate at broader infrastructural and ecological levels, where the individual

researcher has little control. Thus, it is often difficult to determine responsibility or place blame, as this can occur within and across multiple agential dimensions.

Within this arena, one can ask questions such as: How might tech design or implementation impact users in unanticipated ways; e.g., if an operating system’s latest update is so complicated that it makes users feel stupid; when users pay extra for female characters in games but male avatars are free? How might the design of one system impact other systems or in unexpected ways? For example, if a search engine result displays a person’s Twitter contributions to a particular hashtag, their data are exposed in a different medium than intended. It is common that data or algorithms collected or designed for one audience or purpose are analyzed or used by others without permission. What have been some of the consequences? How might the presentation of research in scientific or public sphere impact people; e.g., when data or images shown in presentations reveal personal information that is used in other contexts; or when tentative findings are interpreted out of proportion?

Impact Arena 3. Use of Data

This arena builds on the strengths of feminist, critical, and cultural studies domains, where the processes and politics of categorization, marginalization, and power are central concerns. Beyond data gathering and management, this category focuses on how data are used by various stakeholders to do other things. Data analytics build inferences or generalizations about people or communities, which can reinforce or resist social classification and categories. These generalizations are used as grounds for building specific technologies, creating particular structures, policies, and laws.

Here, we can ask questions such as: How might data aggregation or multi-source analytics bypass basic rights? For example, data analytics in policing means “daily activities, now codified as data, can be marshaled as evidence *ex post facto*,” whereby the datasets as well as the analytic processes obscure the origin points of police investigations and essentially build predictive systems of policing that bypass (or strip) citizen protections against many forms of warrantless surveillance (Brayne, 2017). What is highlighted and hidden when relying on data rather than other forms of evidence? For example, when knowledge about people relies on information collectable as data, it devalues non-datafiable characteristics and aspects. This concern is also evident when people who are not connected or do not produce much data are left out of studies, or when data are given higher truth value than human accounts in courtrooms.

Impact Arena 4. Future Making

This arena builds on the strength of speculative domains, which focus on possible trajectories, often using extreme dystopian or utopian perspectives to vivify how these futures

are experienced. This arena of impact assumes that all research, design, and development will have impact on the shape of future social, political, or cultural formations. This future making includes even banal actions taken at the micro level by individual researchers. Focusing on the idea that we are in the process of making possible certain futures through our own decisions and actions, whether minimal or monumental, can enable us to ask questions that speculate about both positive and negative transformations.

What might be the largescale impact if certain technological development trends became permanent fixtures in society (e.g., relegating key decision making to AI systems; using personalization algorithms to polarize newsfeeds)? How might data or analytical tools used in one context shift to other contexts to exert or reinforce structures and practices of privilege and marginalization (e.g., when genetic data are used to predetermine abilities; when health companies use quantified self-tracking data to assess insurance rates; when automated profiling puts people into stereotyped categories for targeted news, services, and marketing)? How might device or platform designs shape future wellbeing (e.g., the rise of public shaming facilitated by design of twitter; shifting cultural notions of how we think about friendship; increase in physical problems resulting from sitting, typing, and looking down)?

This model sits alongside those that emphasize the importance of asking tough, even unanswerable, questions. This is the core activity of a proactive ethics stance, since it recognizes that ethics are never in a fixed state. We recognize this when we apply ethics to our everyday lives, but in research and design situations, regulatory frameworks encourage a false fixity. Although implemented with good intentions, strict procedure-driven or even concept-driven ethics guidelines can actually undermine strong decision making by limiting it to sets of checkboxes to be ticked off, hoops that need to be jumped through, or bureaucratic tangles that need to be strategically avoided. When decisions get made by individuals, chains of consequences are enacted. Focusing on various scales and arenas of impact enables an anticipatory logic that works with the complexity of causality and accepts that impact is inevitable.

The future-looking-backward orientation of an impact model serves a regulatory function, in that the anticipatory analysis modulates decision making. Rather than the regulation being asserted from the outside—by institutional or historical contexts, it is derived continuously and iteratively. This more internal form of regulation can be more actively attuned to the situation, as it is an ongoing activity. This not only transforms the regulatory function from the noun to verb form (regulation to regulate) but also distributes agency for ethics, or “doing the right thing,” in a more balanced way.

Conclusion

This text is intended as one of the many resources to aid researchers and designers in building strong ethical platforms to undergird our inquiry practices as we engage in social

research, technology design, and policy making. Within the *Social Media + Society* special issue of Ethics as Method, it functions as a primer for the contemporary stance of contextual ethics, to provide a cursory sketch of key issues and shifts in larger debates. It aims to aid readers who are interested, but not necessarily immersed in, these debates understand the starting point for most of the specific contributions. In the broader context of everyday research and development, it is intended as a resource and a tool to think with.

Despite its critical stance toward error-avoidance and concept-driven models ethics, this piece does not mean to undermine the good work of ethics oversight mechanisms by pitting committees and regulators as evil counterparts to context-sensitive approaches. Rather, the aim has been to emphasize the importance of taking the ethics discussion to a level that allows researchers, designers, and policymakers to engage with ethics for the sake of a better outcome, not a streamlined process. The particular argument in this article urges readers to look at ethics as moment-by-moment choices that naturally have consequence—on people, be they defined as human subjects, digital signals, users, profiles, or data points; on the (re)production of particular social orders, be it through forms of classification, categorization, or development of structures and policies; and on our collective socio-political, cultural, economical, and environmental futures.

Finally, the focus on impact invites renewed attention to the locus of responsibility as not just a legal placement but a moral one. Despite, or perhaps because of a widespread dissipation of responsibility in the era of networked flows of information (Hilty et al., 2005, p. 265), whereby it could be argued that nobody takes responsibility because the distribution of cause and effect is too huge to be gathered, it has become necessary for researchers, computer scientists, engineers, and other knowledge and technology designers to bear this burden alongside regulators. Roger Silverstone (2007), in making a similar argument for journalists, draws on Hans Jonas (1984) to distinguish between formal and substantive responsibility. The former is when one takes responsibility for their own actions, the latter a responsibility for the condition of the other. Embracing the inevitable impact of one's interference in the world is one part of taking a stance of substantive responsibility. This reflects some of the same assumptions undergirding a feminist ethic of care, a relational approach with all the concreteness, partiality, difficulty, and situatedness that goes along with care for others.¹³ Of course, morality requires agency, and understanding the limits dictated by extant legalities and regulatory procedures is only the first of our obligations; the rest emerge as we move through and make decisions in larger ecologies, recognizing and then working alongside multiple human and non-human stakeholders to create better possible futures.

Acknowledgements

I am particularly thankful to Eva Svedmark, whose thoughtful development of work in this area has helped me come to richer understanding of these issues.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. While there are too many to list here, including the pieces in the special issue within which the present article appears, one might begin with some edited collections, including Zimmer & Kinder-Kurlanda (2017), Markham & Buchanan (2012), Ess & the AoIR Ethics Working Committee (2002), as well as the ongoing work of such groups as the Council for Big Data Ethics (<https://bdes.datasociety.net/>), the Center for Information Policy Research (<http://cipr.uwm.edu/>), the Center for Digital Ethics and Policy (<http://www.digitalethics.org/>), and the AoIR Ethics Working Committee (<http://aoir.org/ethics/>).
2. For example, the Helsinki Report and the Belmont Report, statements ratified in multiple countries, emerged as responses to ethical breaches in biomedical research, particularly the Nazi medical experiments and the Tuskegee syphilis studies (<http://www.cdc.gov/tuskegee/timeline.htm>). Other well known problematic studies such as the Stanley Milgram obedience experiments (https://en.wikipedia.org/wiki/Milgram_experiment) and the Stanford University prison experiment (<http://www.prisonexp.org/>) have been important in developing codes of ethics around social psychological research. Key basic tenets shared by these statements include the fundamental rights of human dignity and autonomy, minimization of harms, alongside the fundamental responsibility of research to operate with integrity, respect, justice, and beneficence.
3. This is only one of several explanations. See other excellent responses by Zimmer (2014), boyd (2014b), and Grimmelman (2015).
4. See for example McNeal, 2014; Tufekci, 2014
5. Charles Ess has a number of important essays that build this concept in the context of the digital. (Ess, 1996; 2013). See also the related work on the similar concept of casuistic, or case-based reasoning by Heidi McKee & James Porter (2009).
6. I am grateful to Eva Svedmark for providing these cases. She has since written extensively about them in her doctoral work (2017).
7. While Storm King mentioned this concept quite early in the history of the public internet, the concept has been developed considerably since. See for example Bromseth, 2003; McKee & Porter, 2009; Sveningsson, 2004; Whiteman, 2012; Zimmer, 2010).
8. This is a gross understatement for those who would classify Kirkegaard's actions as illegal, wrong, or otherwise egregiously defying international standards for ethical behavior of scientists.
9. <https://www.engadget.com/2016/05/17/publicly-released-okcupid-profiles-taken-down-dmca-claim/>
10. <http://fortune.com/2016/05/18/okcupid-data-research/>
11. <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html>

12. See also Hilty, et al., review (2005) of ethical discussions among technologists to highlight how distance from the data can obscure the potential outcomes or impact of design and deployment.
13. This stance emerged in various works of feminist such as Sara Ruddick (1980), Carol Gilligan (1982), Nel Noddings (1984). For overviews, see Preissle & Han in the Handbook of Feminist Research (2012) or the entry on feminist ethics in the Stanford Encyclopedia of Philosophy (Tong & Williams, 2009).

ORCID iD

Annette N. Markham  <https://orcid.org/0000-0001-8152-2473>

References

- boyd, d. (2014a). *It's complicated: The social lives of networked teens*. New Haven, CT: Yale University Press.
- boyd, d. (2014b). What does the Facebook experiment teach us? Growing anxiety about data manipulation. *Medium.com*. Retrieved from <https://medium.com/message/what-does-the-facebook-experiment-teach-us-c858c08e287f>
- Bromseth, J. (2003). Ethical and methodological challenges in research on net-mediated communication. In M. Thorseth (Ed.), *Applied ethics in Internet research* (pp. 67–86). Trondheim, Norway: NTNU University Press.
- Ess, C. (1996). The political computer: Democracy, CMC, and Habermas. In C. Ess (Ed.), *Philosophical perspectives on computer-mediated communication* (pp. 197–230). Albany: SUNY Press.
- Ess, C. (2013). *Digital media ethics*. London, England: Polity Press.
- Ess, C., & the AoIR Ethics Working Committee. (2002). *Ethical decision-making and Internet research: Recommendations from the AoIR Ethics Working Committee*. Chicago, IL: Association of Internet Researchers (AoIR).
- Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.
- Grimmelmann, J. (2015). *The law and ethics of experiments on social media users*. Unpublished manuscript. Retrieved from <http://ctlj.colorado.edu/wp-content/uploads/2015/08/Grimmelmann-final.pdf>
- Harmon, A. (2010, April 21). Indian tribe wins fight to limit research of its DNA. *New York Times*. Retrieved from <https://www.nytimes.com/2010/04/22/us/22dna.html>
- Hilty, L., Behrendt, S., Binswanger, M., Bruinink, B., Erdmann, L., Fröhlich, J., . . . Würtenberger, F. (2005). *The precautionary principle in the information society: Effects of pervasive computing on health and environment*. Bern, Switzerland: TA-SWISS.
- Jonas, H. (1984). *The imperative of responsibility: In search of an ethics for the technological age*. Chicago, IL: University of Chicago Press.
- King, S. (1996). Researching Internet communities: Proposed ethical guidelines for the reporting of results. *The Information Society*, 12(2), 119–128. doi:10.1080/713856145
- Markham, A. (2012). Fabrication as ethical practice: Qualitative inquiry in ambiguous Internet contexts. *Information, Communication, & Society*, 15, 334–353.
- Markham, A. (2015, May 8). *Assessing the “creepy” factor: Shifting from regulatory ethics models to more proac-*

- tive approaches to “doing the right thing” in technology research. Talk at Microsoft Research Labs, Cambridge, MA. Retrieved from <https://www.microsoft.com/en-us/research/video/assessing-the-creepy-factor-shifting-from-regulatory-ethics-models-to-more-proactive-approaches-to-doing-the-right-thing-in-technology-research/>
- Markham, A., & Buchanan, E. (2012). *Ethical decision-making and Internet research: Version 2.0*. Chicago, IL: Association of Internet Researchers (AoIR).
- Marwick, A., & boyd, d. (2014). Networked privacy: How teens negotiate context in social media. *New Media & Society*, 16(7), 1051–1067. doi:10.1177/1461444814543995
- McKee, H. A., & Porter, J. E. (2009). *The ethics of internet research: A rhetorical, case-based process*. New York, NY: Peter Lang.
- McNeal, G. (2014, June 28). Facebook manipulated user news feeds to create emotional responses. *Forbes*. Retrieved from <https://www.forbes.com/sites/gregorymcneal/2014/06/28/facebook-manipulated-user-news-feeds-to-create-emotional-contagion/#597996aa39dc>
- Metcalf, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society*, 3(1), 1–14. doi:10.1177/2053951716650211
- Moor, J. (1985). What is computer ethics? *Metaphilosophy*, 16, 266–275.
- Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Palo Alto, CA: Stanford University Press.
- Noddings, N. (1982). *Caring: A feminine approach to ethics and moral education*. Berkeley: University of California Press.
- Preissle, J., & Han, Y. (2012). Feminist research ethics. In S. Hesse-Biber (Ed.), *Handbook of feminist research: Theory and praxis* (pp. 583–605). Thousand Oaks, CA: SAGE.
- Ruddick, S. (1980). Maternal thinking. *Feminist Studies*, 6, 342–367.
- Silverstone, R. (2007). *Media and morality*. Cambridge, UK: Polity.
- Sterling, R. (2011). Genetic research among the Havasupai—A cautionary tale. *AMA Journal of Ethics*, 13, 113–117.
- Svedmark, E. (2016). *Becoming together and apart: Technoemotions and other posthuman entanglements* (Doctoral dissertation). Umeå, Sweden: Umeå University. Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-120195>
- Svedmark, E., & Nyberg, A. (2009). Om det privata i publika och digitala rum [About privacy in public and digital spaces]. In M. Nyman (Ed.), *Semig: Unga Om Sex Och Internet* [See me: Youth about Sex and the Internet] (pp. 354–383). Stockholm, Sweden: Davidsons Tryckeri AB.
- Sveningsson, M. (2004). Ethics in Internet ethnography. In E. A. Buchanan *Readings in virtual research ethics: Issues and controversies* (pp. 45–61). Hershey, PA: Information Science Publishing.
- Tong, R., & Williams, N. (2009). Feminist ethics. In *The Stanford encyclopedia of philosophy* (Winter 2016 edition). Retrieved from <https://plato.stanford.edu/archives/win2016/entries/feminism-ethics/>
- Tufekci, Z. (2014). Facebook and engineering the public: It’s not what’s published (or not), but what’s done. *Medium.com*. Retrieved from <https://medium.com/message/engineering-the-public-289c91390225>
- Whiteman, N. (2012). *Undoing Ethics: Rethinking practice in online research*. New York, NY: Springer.
- Zimmer, M. (2010). “But the data is already public”: On the ethics of research in Facebook. *Ethics and Information Technology*, 12, 313–325.
- Zimmer, M. (2016, June 3). OKCupid and the ethics of big data. *Ampersand*. Retrieved from <http://blog.primr.org/okcupid-ethics-big-data-research/>
- Zimmer, M., & Kinder-Kurlanda, K. (Eds.). (2017). *Internet research ethics for the social age: New challenges, cases, and contexts*. New York, NY: Peter Lang.
- Zook, M., Barocas, S., boyd, d., Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., & . . . Pasquale, F. (2017). Ten simple rules for responsible big data research. *PLoS Computational Biology*, 13(3), e1005399. Retrieved from <https://doi.org/10.1371/journal.pcbi.1005399>

Author Biography

Annette N. Markham is Professor MSO of Information Studies at Aarhus University, Denmark, and Affiliate Professor of Digital Ethics in the School of Communication at Loyola University, Chicago. Annette is internationally recognized for developing epistemological frameworks for rethinking ethics and qualitative methods for digitally saturated social contexts. A long-time member of the digital research community, Annette conducts sociological and ethnographic studies of how identity, relationships, and cultural formations are constructed in and influenced by digitally saturated socio-technical contexts. For more information, please see annetmarkham.com