Original Research Article

Where are human subjects in Big Data research? The emerging ethics divide

Jacob Metcalf[1] and Kate Crawford[2]

Abstract

数据科学的研究实践与现有研究伦理工具之间的断裂正在不断加剧

There are growing discontinuities between the research practices of data science and established tools of research ethics regulation. Some of the core commitments of existing research ethics regulations, such as the distinction between research and practice, cannot be cleanly exported from biomedical research to data science research. Such discontinuities have led some data science practitioners and researchers to move toward rejecting ethics regulations outright. These shifts occur at the same time as a proposal for major revisions to the Common Rule—the primary regulation governing human-subjects research in the USA—is under consideration for the first time in decades. We contextualize these revisions in long-running complaints about regulation of social science research and argue data science should be understood as continuous with social sciences in this regard. The proposed regulations are more flexible and scalable to the methods of non-biomedical research, yet problematically largely exclude data science methods from human-subjects regulation, particularly uses of public datasets. The ethical frameworks for Big Data research are highly contested and in flux, and the potential harms of data science research are unpredictable. We examine several contentious cases of research harms in data science, including the 2014 Facebook emotional contagion study and the 2016 use of geographical data techniques to identify the pseudonymous artist Banksy. To address disputes about application of human-subjects research ethics in data science, critical data studies should offer a historically nuanced theory of "data subjectivity" responsive to the epistemic methods, harms and benefits of data science and commerce.

Keywords

Data ethics, human subjects, Common Rule, critical data studies, Big Data

公开可用的数据可能被用于广泛的二次用途，包括与其他数据集结合使用，从而对个体和社区造成严重风险

Introduction

Critical data studies is in its infancy, but it faces a substantial challenge: as the practice of data science surges ahead, we lack a strong and rigorous sense of ethical parameters for scientific research. There are several problems emerging. First, there is a growing divide between established systems of research ethics in more traditional disciplines and the dynamic norms and research methods of Big Data. Big Data research methods exacerbate a long-standing tension between the social sciences and research regulations that are geared to the methods and harms of biomedical research. Second, US research regulations (both the current rules and proposed revisions) exempt projects that make use of already existing, publicly available datasets on the assumption that they pose only minimal risks to the human subjects they document.[1] But this assumption is founded on a misconception. Publicly available data can be put to a wide range of secondary uses, including being combined with other data sets, that can pose serious risks to individuals and communities. This is one of several risks that are being overlooked in the current debates about the ethics of Big Data studies.

[1]Data & Society Research Institute, New York, NY, USA
[2]Microsoft Research, MIT Center for Civic Media, New York University Information Law Institute, New York, NY, USA

Corresponding author:
Jacob Metcalf, Data & Society Research Institute, 36 W 20th St, New York, NY 10011, USA.
Email: jake.metcalf@datasociety.net

For example, in 2016, a group of researchers published a study that sought to reveal the identity of British artist Banksy, who has sought to keep his real name out of the public domain (Hauge et al., 2016). They used geographical profiling, a technique of statistical inference traditionally used in serial crimes like rape and murder, to hone in to a suspected person. They analyzed the spatial patterns of Banksy's artworks around London and Bristol, and then tracked a particular individual who had been named by the *Daily Mail* as likely to be Banksy.[2] They searched the electoral rolls for this person's former addresses as well as those of his wife, and places where he likely went to school and played football. Then Banksy's public artworks were mapped against these streets and neighborhoods. They investigated no other "suspects" but argue that their findings support those of the *Daily Mail*. The researchers claim that their approach could be useful for early identification of terrorists, as "terrorists often also engage in low level activities such as vandalism, graffiti, anti-government leaflet distribution, and banner posting" (Hauge et al., 2016: 5).

There are many questions that could be asked of this study, not least about the correlation between graffiti and terrorism. But for our purposes, we will only focus on the "ethical note" that appeared at the end of the article: "the authors are aware of, and respectful of, the privacy of [subject name removed] and his relatives and have thus only used data in the public domain" (Hauge et al., 2016: 5). This claim is particularly striking, as it is difficult to see how tracking a specific individual (and their family) to such an invasive degree could be considered respectful of their privacy. But there are now so many data sets about individuals in the public domain, that, while relatively innocuous in themselves, become highly identifying when brought together. The Banksy study is not a large-scale data study, but it echoes the argument made by many Big Data researchers that they are absolved of ethical concerns by pointing to the "publicness" of the data they use. By applying specialized tools for tracking terrorists, Hauge et al. revealed sensitive patterns of movement over several decades. Though they only delved into public data stores, they exploited everything they could find about an artist's personal (and creative) life, and cross-referenced it with the details of a private citizen, in order to expose an identity that the artist sought to keep secret.

The researchers who published the Banksy study say they went through review from an independent ethics board, and while we cannot see their determination, it is likely that they were allowed to track their suspected individual because the data was public as that is a common standard across research ethics regulations.[3] We argue it is a useful case study of why public data can be incredibly invasive, and potentially harmful.

Critical data studies has an important role to play in analyzing and clarifying these issues by situating questions of data ethics regulations and norms within a historical and discursive analysis of the core concepts and norms of research ethics in general. By historicizing extant research ethics norms and regulations, we are able to see the disjunctions with the epistemic conditions of data sciences as one more site of negotiation and improvement rather than an implacable conflict.

Big Data stretches our concepts of ethical research in significant ways (boyd and Crawford, 2012). It moves ethical inquiry away from traditional harms such as physical pain or a shortened lifespan to less tangible concepts such as information privacy impact and data discrimination. It may involve the traditional concept of a human subject as an individual, or it may affect a much wider distributed grouping or classification of people. It fundamentally changes our understanding of research data to be (at least in theory) infinitely connectable, indefinitely repurposable, continuously updatable and easily removed from the context of collection. By doing so, it forces us to grapple with the ways in which familiar and practical ethical constraints depended upon research data being temporally and contextually constrained and restricted by technical infrastructures and financial cost. Further, data science methods create an abstract relationship between researchers and subjects, where work is being done at a distant remove from the communities most concerned, and where consent often amounts to an unread terms of service or a vague privacy policy. Together, these shifts are hard to quantify and ameliorate (Zwitter, 2014), frustrating the familiar ethical practices outside of biomedical research. So while extant research ethics and regulations are far from a perfect fit for the methods of Big Data, there is real urgency to define what a "human subject" is in Big Data research and critically interrogate what is owed to "data subjects." What lessons might we learn from the history and implementation of human-subjects research protections in order to better address these growing conceptual and structural discontinuities? How have other non-biomedical fields of science confronted the question of ethics through a critical lens?

Part of the difficulty here is that the precursor disciplines of data science—computer science, applied mathematics and statistics—have not historically considered themselves as conducting human-subjects research. Even though statistics do ultimately represent people, research into math, computational capacity and other numeric modes of analysis rarely exhibited the types of human subjects concerns that are baked into research ethics regulations designed to handle the types of harms found in biomedical research. Such regulatory definitions rest on a set of ethical and epistemic

大数据在很大程度上扩展了我们的伦理研究概念（boyd和Crawford，2012）。它将道德调查从身体疼痛或寿命缩短等传统伤害转移到信息隐私影响和数据歧视等不太有形的概念上。

assumptions which are now under contestation due to Big Data methods.

For example, data analytics techniques rarely appear as a direct "intervention" in the life or body of an individual human being, which is one of the key requirements for research to be regulated in the USA (Department of Health and Human Services, 2009). The action of Big Data analytics happens mostly at a remove from the point of data collection, which is the most plausible analog for an "intervention." Instead, it is focused on data sets that likely have a long lifespan and may be continuously updated and re-analyzed. Similarly, the Common Rule assumes that data which is already publicly available cannot cause any further harm to an individual.[4] Yet this fails to account for data analytics techniques that can create a composite picture of a person from disparate datasets that may be innocuous on their own but produce deeply personal insights when combined (Crawford and Schultz, 2014). The assumption (codified in law) that individual harm is the only type of risk researchers are required to track and mitigate undercuts the ability to see and account for harms that affect communities or produce "networked harms" (boyd et al., 2014).

Implicitly, the existing ethics regulations promote a historically situated understanding of "research subjectivity" that is clearly eroded by data science. The assumptions about what constitutes an intervention, when and how consent should occur and what types of harms are relevant, all add up to a picture of the human-research subject that is out of step with large-scale data practices. If the familiar human subject is largely invisible or irrelevant to data science, how are we to devise new ethical parameters? Who is the "data subject" in a large-scale data experiment, and what are they owed?

In this paper, we offer a preliminary examination of how critical data studies might generate a theory of data subjectivity that would enable responsible scientific practice with Big Data methods. We map the discontinuities between research regulations and data science, focusing in particular on human-subjects protections and the 30 year debate in the USA about the regulation of human-sciences research. We show that while the proposed revisions to the Common Rule are helpful in terms of making research ethics regulations more flexible and scalable to different research methods and types of risk, they problematically exclude data science wholesale in situations that still present serious risks. These exclusions are based on questionable assumptions about publicly available data, researcher–subject relationships and the very nature of "intervention" into the daily lives of those whose data is held within research databases.

Data science, social science and the complicated human subject

There are a variety of reasons why the predecessors of data science—applied mathematics, statistics and computer science—have had little contact with the infrastructures of ethics review. For the most part, the basic science conducted in these fields has had only distant contact with human data. Researchers represent themselves as dealing with systems and math, not people—human data is treated as a substrate for testing systems, not the object of interest in itself. The infrastructures of human-subjects protections have largely accepted this position, but where Institutional Review Boards (IRBs) have engaged with data science there appears to be mutual confusion. University-based IRBs are overwhelmingly oriented toward the methods common to biomedical and psychological experimentation in which interventions carry clear risks to individual subjects. Now that data science techniques profoundly affect human lives, the computational and mathematical disciplines are in urgent need of strong, adaptable ethical frameworks.

A robust approach to data ethics should interrogate how subjectivity is constructed in research datasets. Critical data studies have routinely demonstrated that it is deeply mistaken to treat research data as neutral and raw (see, for example, Bowker, 2005; Gitelman, 2013). Datasets and algorithms have historical, material specificity that is laden with political and ethical values. As data science moves toward interpreting and manipulating social structures and behaviors, often drawing on the interpretative tools of social science, these values become both more evident and more consequential. Hence, there is a need for more nuanced ethical research processes. We suggest that as computer science is being drawn into a closer orbit with social science we need to re-examine the rocky relationship between the social sciences (and to a lesser extent, the humanities) and research ethics infrastructures. In this closer conversation between the norms of social science research and the emerging practices of data science, there have been no clear conclusions about what counts as a human subject, and little research into what protections they might deserve.

Yet, it may be unnecessary to create an entirely new definition of what counts as a research subject in data science. Instead, we advocate for an approach to research subjectivity that is co-emergent with the conditions of research. From the earliest biomedical research ethics documents and policies, the question of how human-subjects get defined has been contested by scientists, physicians and ethicists (Annas, 1992). These debates revolve around norms of trust between researchers and subjects that run deeper than

regulatory definitions. Situating critical data studies within the ongoing, dynamic debates about human subjects—rather than treating it as an entirely new field with unique problems—can remind data scientists and ethicists that we are engaging with a rapidly changing set of research dynamics that should be addressed in context, rather than solely through regulatory decisions.

Historicizing conflicts over ethics regulations

The current debate about human subjects in data science contains echoes of the history of social scientists contesting regulation of their research. Social and behavioral researchers vociferously contested the first drafts of the Common Rule because it consistently applied the same level of scrutiny to medical experiments on humans as sociologists' interviews of humans. Duster et al. (1979) argued that human-subjects protections intended for vulnerable populations can inadvertently reinforce political disparities that have much worse consequences for those populations. Citing a field study of racial housing discrimination that sought to interview landlords, they point to the risk that requiring consent from all parties in the fashion of biomedical research risks excluding certain methods of justice-oriented research. Decades later, social scientists continue to make claims about the codified norms of research ethics regulations (Shea, 2000). For example, Librett and Perrone (2010) claim that ethnography operates at ethical and epistemic odds with human-subjects protections, and that university IRBs undermine ethnographic knowledge and discipline-specific ethical practices by risking confidentiality.

Social scientists have similarly critiqued the application of human-subjects protections for Internet-based research methods (Keller and Lee, 2003; Walther, 2002). Bassett and O'Riordan (2002) argue that Internet research is about cultural texts, not social spaces, and therefore should be considered closer to history or biography and be exempt from research regulations. Neuhaus and Webmoor (2012) similarly contend that much "massified" social science research should instead adopt a model of "agile ethics": utilizing transparent and publicly available ethical commitments on the part of individual researchers in lieu of contractual-informed consent agreements. Over time, all the critiques outlined above have pointed to the problem of lumping disparate types of research together without respect to gradations of potential risks and benefits in their different research methods.

Similarly, the regulatory agencies are criticized for addressing ethics with a one-size-fits-all approach, and then applying those rules inconsistently across similar cases, which creates unfair burdens on researchers and expensive delays to research projects (Abbott and Grady, 2011; Committee on Revisions et al., 2014; Fost and Levine, 2007; Ledford, 2007; Rhodes et al., 2011; Silberman and Kahn, 2011). This can give the impression that research regulation is fundamentally a matter of outsiders with inscrutable agendas interfering with the important work of advancing science and engineering. Given that narrative, it might be understandable why data-intensive researchers would be deeply skeptical of falling under current research ethics regulations. Yet, there are important lessons for data science to be found in an alternative reading: research ethics regulations can be understood as an imperfect embodiment of norms of trust between researchers and subjects in what is ultimately a system of *self-regulation* by researchers. Rather than fretting about the poor fit between data science and biomedical regulations, data scientists should aim for modeling the norms and practices that would build and sustain the public trust necessary to earn the right of effective self-regulation.

Ethical codes often emerge after a crisis event. The Common Rule developed out of a rule-making process initiated in response to a series of breaches to the public trust, especially those committed by physician-researchers. Following the Nazi-era medical atrocities, the Nuremberg Code (Nuremberg Code, 1949) and the Declaration of Helsinki (World Medical Association, 1964) established ethical norms for human-subjects research, while building on the 1931 Guidelines for Human Experimentation (Ghooi, 2011). The Nuremberg Code codified many of our standard principles of ethical research, including that informed consent is required of all subjects, subjects have a right to withdraw at any time without consequence, research must appropriately balance risk and potential reward, and researchers must be well versed in their discipline and ground human experiments in animal trials.

Importantly, ethics codes also serve a number of functions beyond deterring unethical behavior, including creation of a cohesive community identity, responding to external criticism and—most importantly for our purposes—establishing the moral authority for self-regulation (Frankel, 1989; Gaumnitz and Lere, 2002; Kaptein and Wempe, 1998; Metcalf, 2014). The American Medical Association's code was the first-ever code adopted by a medical professional society, and the contemporary version tightly links ethical integrity and "the profession's authority to self-regulate" (American Medical Association, 2015).

But these codes did not carry the weight of law in the USA until after a series of research scandals in the 1960s and 1970s—most notably, the Tuskegee syphilis experiment. This led to the 1974 National Research

Act, which established the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The National Commission's most consequential output was the Belmont Report in 1979. The Belmont Report was not itself a regulation and explicitly avoided policy recommendations. Instead, it was adopted by the Health and Human Services Administration as the source of core principles for the rule-making procedures that ultimately generated the Common Rule. For most researchers, the most significant aspect of the Common Rule was establishing IRBs, which act as independent panels that review research proposals to assess possible harms to human subjects. Unlike many forms of regulation, the Common Rule invests research institutions with the power and responsibility to self-regulate through these boards: it confers authority and establishes a relationship of trust.

The Belmont Report established three basic principles as foundational for biomedical research governance: respect for persons, beneficence and justice. While these principles get the bulk of attention, perhaps the most consequential contribution for fields like data science is the attempt to define the boundary between research and practice. While acknowledging that the distinction is imperfect in many edge cases, the Report states that this need not cause substantial confusion: "the general rule is that if there is any element of research in an activity, that activity should undergo review for the protection of human subjects" (National Commission for the Protection of Human Subjects, 1979). When routine medical practice veers toward untested territory, it becomes necessary to signify that a matter of clinical care has been "made the object of formal research at an early stage" in order to guarantee its safety (National Commission for the Protection of Human Subjects, 1979).

A vital characteristic of the Belmont Report from the perspective of data science is this close pairing of epistemic and ethical commitments. But the practice/research distinction has led to some absurd outcomes, such as tightly regulating best practices research but not regulating untested changes to practice, and occasionally shutting down best practices research regarding common clinical practices. A notorious case of such overreach resulted from a study of infection-control protocol when inserting catheters during intensive care (Gawande, 2007; Kass et al., 2008; Pronovost et al., 2006; Thompson et al., 2012). The study showed that requiring physicians to follow a simple procedural checklist of commonly accepted practices saved 1500 lives and $200 million in just 18 months. But the research team faced penalties for crossing the codified practice/research distinction as interpreted by federal agencies and not getting informed consent from

each patient and practitioner (Office for Human Research Protections, 2008b).

Tom Beauchamp, one of the staff philosophers for the Belmont Report, has recently suggested that the practice/research distinction will grow increasingly complicated due to intensive data collection (Beauchamp 2011a, 2011b; Beauchamp and Saghai, 2012). Indeed, scientific and technological advances will periodically alter the research/practice topology and it would be a mistake to rest on the distinction as the guarantor of the difference between ethical and unethical activities.[5] We argue that there has been a lack of attention to the *social roles* that are codified in the research/practice distinction. The physician–patient relationship is a largely unique social relationship in which the physician is invested with tremendous trust to make decisions in the best interest of the patient. Regulations built around the research/practice distinction can be read as a method for signaling and negotiating temporary changes to that relationship—a patient must be informed, and consent to, situations in which a physician may no longer be making or be able to make decisions in the best interest of the patient. In a research context, a physician has the best interest of the social collective as an explicit competing interest to the well-being of the patient. In the long arc, research methods, epistemic commitments and ethical-social obligations are deeply interconnected.

Yet, there is no easy analogue for the physician-researcher in data science (or, for that matter, in many other fields). And the iterative nature of algorithmically driven data analytics blurs the line between research and practice. Thus, there is no easy route to use the research/practice distinction as a trigger for ethical review data science in the fashion of the Belmont Report—instead we need substantive, critical and nuanced assessments of ethics regulations. That assessment begins with the understanding that research ethics regulations are an imperfect codification of the hard-won, often-contested and evolving social trust invested in practitioners and researchers. Importantly, the ethics regulations targeted by critics (Meyer, 2014, 2015), and the codes that informed those regulations, have played no small part in maintaining that trust over time. Insofar as physician–researchers contributed to the formation of those codes and regulations, and the broader research community assented to them (even if begrudgingly), research ethics regulations have built the bedrock of trust that has ultimately enabled research to occur at all. Therefore, even if the research/practice distinction as codified in the Common Rule proves too unwieldy for the methods of data science, we still need regulatory options that build trust between data practitioners and data subjects (Polonetsky et al.,

2015). Namely, what are the actionable ethical obligations data scientists and practitioners have for the well-being of data subjects? How do we assess that those obligations are being met? Answering these questions is essential to developing a trustworthy system for data science experiments that can influence the future of millions of people.

A new Common Rule? The implications for data science

The research ethics challenges posed by data science are unfolding just as the US Department of Health and Human Services is proposing the first major revisions to the Common Rule in over two decades. In September 2015, the HHS released a Notice of Proposed Rule Making (NPRM), which is essentially a first draft of revisions that may eventually have the force of law (Department of Health and Human Services, 2015a, 2015b). Many of the areas which they cover have substantial bearing on data-intensive research techniques,[6] opening up a rare moment of major regulatory change just as conversations about data ethics are becoming prominent. Broadly speaking, the NPRM creates a greater range of regulatory categories that are meant to be indexed to empirical measurements of risk. The NPRM provides much more specific guidance to IRBs for determining the proper level of oversight for research projects, reducing bureaucratic burdens, clarifying the status of biobanked specimens and streamlining the informed consent process for subjects and researchers alike. But one unintended consequence may be that many data-intensive projects will be permanently outside any ethics review process whatsoever. If it is decided that there is no "human-subjects research" being done in data science, we argue this will be perilous for the subjects of Big Data studies, as well as for the nascent trust in the field.

The proposed revisions note that the relationships between subjects, researchers and data are shifting around multiple poles simultaneously: subjects care more about managing their data, the risk profile of human-subjects data is changing unpredictably and researchers are increasingly able to access data without interacting with the subjects (Department of Health and Human Services, 2015b: 30–31). Beyond these questions, we think critical data studies should also consider the power asymmetries of large-scale data studies, and the shifting concepts of consent intervention and agency. For the purposes of this article, we will focus on what we see as the most consequential change for data science: the major growth of categories of research that receive little or no oversight from IRBs on the problematic premise that publicly available data poses minimal risk to human subjects.[7] By tracking how these proposed

ethics regulations fail to address the sorts of harms involved in data science, we illustrate why a theory of data subjectivity is needed in critical data studies.

The NPRM document acknowledges that technology is rapidly altering the epistemic conditions and risk profiles of human-subjects research:

> The sheer volume of data that can be generated in research, the ease with which it can be shared, and the ways in which it can be used to identify individuals were simply not possible, or even imaginable, when the Common Rule was first adopted. (Department of Health and Human Services, 2015b: 29)

The NPRM also notes that effectively scaled regulation requires empirical measurement of risks, particularly with regards to defining minimal risk (see, for instance, Department of Health and Human Services, 2015b: 236, 433–436). Yet, as we will show, the proposed changes would make it highly unlikely that IRBs could track or ameliorate those risks. How the revised Common Rule will address data science methods outside of biomedicine largely depends on some critical regulatory definitions. In our interpretation, data science appears to be largely excluded from oversight by the definition of human-subjects research, *excluding* all research that uses publicly available datasets and *exempting* (minimal oversight) research involving secondary use of identifiable data acquired for non-research purposes.

The NPRM's definition of human subjects results in IRBs being tasked with reviewing any research that risks placing private, identifiable (and some types of re-identifiable) data about individuals into public hands, or anything that requires an interaction or intervention in the subject's life to obtain that data (Figure 1). The NPRM's definitions result in some linguistically odd outcomes because some activities that are clearly research and clearly about humans fall outside of it, particularly if the methods generate sufficient distance between the researcher and the subject. Significantly, data science often falls into that odd linguistic gap: research about humans that is not human-subjects research. Ioannidis (2013) calls this the "oxymoron of research that is not research," when research is considered simultaneously powerfully insightful about human lives, but inconsequential when accounting for potential harms. Data science researchers are often able to gain access to highly sensitive data about human subjects without ever *intervening* in the lives of those subjects to obtain it. They may predict, or infer it, or gather it from disconnected public data sets. Here too, we think critical data studies has much work to do in determining what constitutes an "intervention" in the lives of data subjects. For example, are
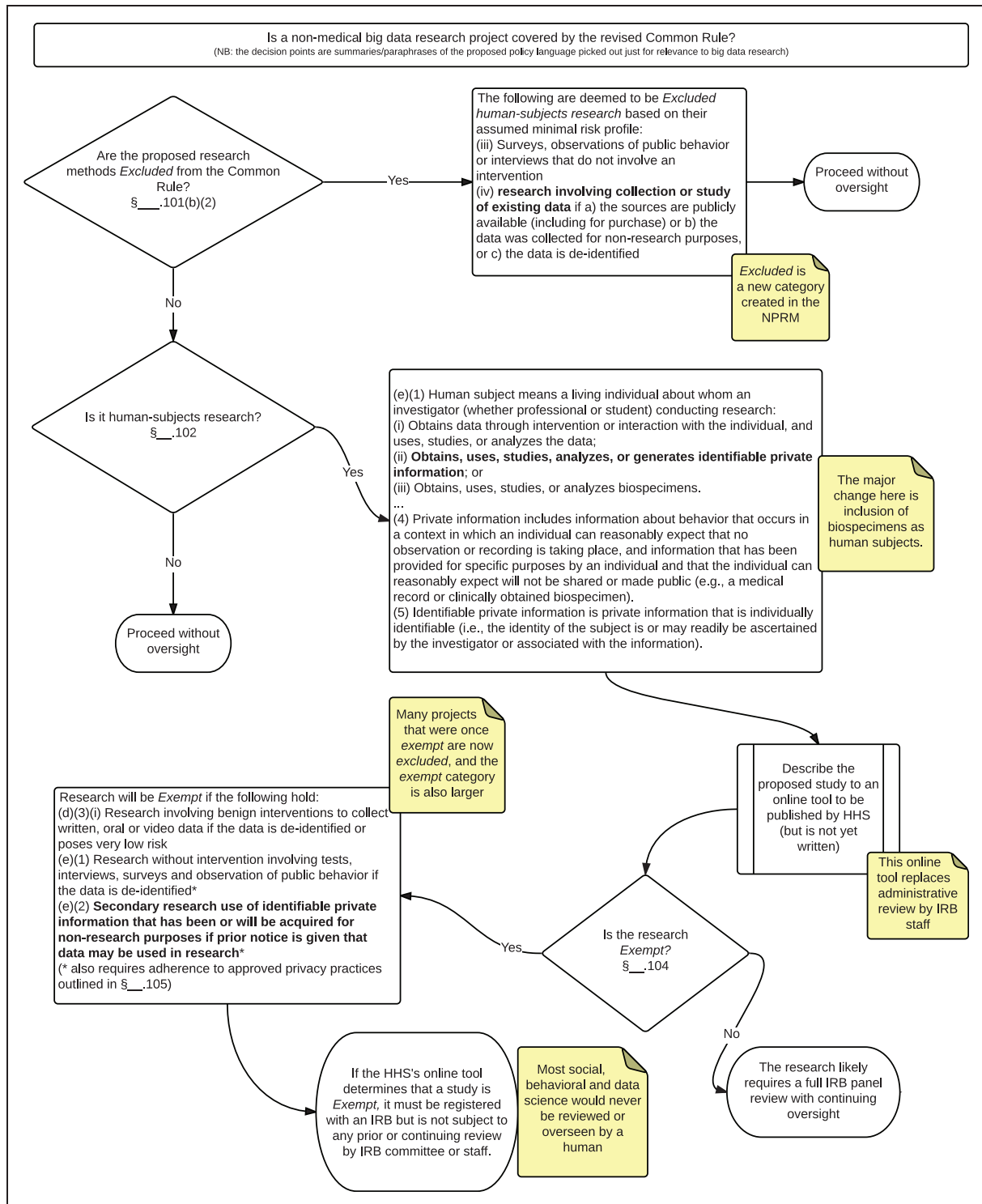
Figure 1. Decision tree for determining whether data science research is covered by the Common Rule as human-subjects research.

predictions "interventions"? Should connecting previously separate records from multiple public databases be considered creating a type of new data?

The criteria for human-subjects protections depend on an unstated assumption that we argue is fundamentally problematic: that the risk to research subjects depends on *what kind* of data is obtained and *how* it is obtained, not *what is done* with the data after it is obtained. This assumption is based on the idea that data which is public poses no new

risks for human subjects, and this claim is threaded throughout the NPRM. While this may have once been a reasonable principle, current data science methods make this a faulty assumption. As data science drives significant changes to *how* we know by creating new knowledge through tying together previously disconnected datasets (Kitchin, 2014; Mayer-Schönberger and Cuvier, 2013), we should expect the ethical consequences of *what* we know to also become significantly knottier (Jackson et al., 2014; Polonetsky et al., 2015). Indeed, the very premise of Big Data analytics is that we can repeatedly generate new, unanticipated knowledge out of already existing measurements, yet the Common Rule revisions would a priori exclude the possibility that it could pose new risks to individuals.

The NPRM responds to complaints of lumping together social science and biomedical research in a one-size-fits-all schema by proposing to adopt oversight scaled to empirical measurements of risk (Department of Health and Human Services, 2015b: 22).[8] But practically speaking, the proposed solution results in far fewer non-biomedical projects passing through the hands of IRBs. This may have the unintentional effect of removing new and emerging categories of risk from review.

For example, the NPRM proposes to newly / exclude/ (meaning no review) "certain research activities that are sufficiently low risk and nonintrusive that the protections provided by the regulations are an unnecessary use of time and resources, whereas the potential benefits of the research are substantial" (Department of Health and Human Services, 2015b: 65).[9] Currently, the Common Rule *exempts* (meaning minimal review) research that makes use of existing data, documents, records and specimens if that data is publicly available (including for purchase) or was recorded by researchers in a way that cannot be used to identify the subjects.[10] The NPRM proposes to *exclude* (meaning no review) such research prior to any review because the publicness of the data means it should pose no new risks to subjects and the researchers have no direct interaction with the subjects. So long as the data is public, the investigator does not identify or contact the subjects, and the investigator does not re-identify the subjects, then they are excluded from ethics review (Department of Health and Human Services, 2015b: 90). Simply put, this is a strong move toward excluding all research using public datasets from ethics regulation.[11]

In effect, the definitions of exempt and excluded research in the NPRM mean that most non-medical data science will receive very little review. The proposed changes will include privacy safeguards in the form of best practices for protecting sensitive data, which IRBs can use as a list of acceptable practices. Those privacy safeguards are not yet written.

Taken together, the revisions mean that research which re-uses de-identified or publicly available data will largely be excused from ethics oversight as long as it meets unspecified privacy safeguards. Given their definition of human-subjects research, nearly all non-biomedical research would receive *at most* perfunctory oversight due to the assumption that there is little or no risk of harm.

Although publicness of datasets may have once been an adequate proxy for risk, it is no longer an empirically sound assumption. The value-added activities in data science and commerce come from pulling together disparate databases to produce new insights. These experiments often use data that may not appear to be personally identifying, but can become so in combination, generating "predictive privacy harms" (Crawford and Schultz, 2014). The range of harms made possible by data analytics extremely hard to foresee and delimit (Andrejevic, 2014; Michael and Miller, 2013; Nissenbaum 2009, 2011; Polonetsky et al., 2015). The same "publicly available" database that meets the proposed *excluded* criteria may have radically different consequences for a subject when multiple public databases are analyzed together, rendering common privacy and anonymization safeguards insufficient (Barocas and Nissenbaum, 2014). How terms of service define "public" can be very different from how actual human subjects conduct publicness in practice, which complicates the computational measures and personal efforts required to protect privacy (Brunton and Nissenbaum, 2015; Dwork, 2011; Dwork and Mulligan, 2013). For example, consider a research project that would correlate an individual's multiple social media feeds and run a linguistic/semiotic analysis that could reveal potentially damaging information—such as political views, sexual orientation, immigration status and so on (Kosinski et al., 2013). Yet, such a project would appear to pass the NPRM's qualifications for non-review based on *how* and *where* data is collected.

Data science risks falling into a regulatory gap that could undermine public trust. This gap is created by a binary conception of datasets as either public or private, rather than dynamic, networked and readily repurposed. *Publicly available datasets containing private data* describes many of the sources most interesting to data researchers and practitioners, and are arguably most risky for subjects, yet are a priori excluded from any review under the NPRM. We see this as a serious problem, and one that requires a deeper critical analysis before it is encoded into ethics review processes. Data subjectivity creates a more fluid relation to publicness than the familiar models of human subjectivity in existing research ethics

在多个公共数据库结合分析时可能对个体产生完全不同的后果，使常规的隐私保护和匿名化措施变得不足

数据科学风险可能会落入一个监管空白，从而削弱公众信任。这个空白是由于将数据集简单地二分为"公开"或"私有"的观念造成的，而没有考虑数据是动态的、网络化的和可轻松重复的性质

regulations. When datasets about humans become dynamic, flexible and interconnected, then our conceptions of what is owed to data subjects should also be flexible and highly attuned to the specifics of individual cases.

Cases of research harms to data subjects

There have been several recent cases where de-identified data that was released publicly was able to be re-identified, or where data that was assumed to have no identifying features could be correlated with specific populations. For example, in 2013, the New York City Taxi & Limousine Commission released a dataset of 173 million individual cab rides, and it included the pickup and drop-off times, locations, fare and tip amounts. The taxi drivers' medallion numbers were anonymized (hashed), but this was quickly de-anonymized—revealing sensitive information such as any driver's annual income and enabling researchers to infer their home address (Franceschi-Bicchierai, 2015). A data scientist at Neustar Research showed that by combining this data set with other forms of public information like celebrity blogs you could track well-known actors, and predict likely home addresses of people who frequented strip clubs (Tockar, 2014). Another researcher demonstrated how the taxi dataset could be used to speculate which taxi drivers were devout Muslims by observing which drivers stopped at Muslim prayer times (Franceschi-Bicchierai, 2015). From one seemingly innocuous and anonymized data set came many unexpected and highly personal forms of information.

The taxi dataset is arguably a case of open data gone wrong—had the dataset been hashed properly it may have been much harder to de-anonymize. However, other research that makes use of *publicly available private data* from multiple databases has been used to make potentially risk-laden correlations. Danyllo et al. (2013) correlated a dataset from a financial institution with Twitter profiles from a geographic region of Brazil. They were able to produce a social network graph demonstrating that social and geographical relationships cluster around similar levels of credit access. The authors note that this research can be used by financial institutions to gauge credit worthiness based on one's social relationships.

Of course, a case can be made that academic researchers should have access to public datasets in order to fully understand their potential and risk. Furthermore, we would caution against presuming that the worst case uses are inevitable with new forms of knowledge—the same research that could be used to discriminate against credit seekers could be used to

track and ameliorate that discrimination. However, we find it concerning that we know so little about the data subjects in these studies and their expectations about how their private data is used in research. Should Twitter users now expect that their social media activities could affect their ability to get a loan? Is it reasonable to assume that social behavior on Twitter is the same as social relationship outside of Twitter, or is this a spurious correlations that might cause economic harm to particular individuals and communities? If human-subjects research regulations assume that public datasets are inherently harmless, it will be nearly impossible to review the material consequences to the affected data subjects.

These cases and the Banksy data tracking study (Hauge et al., 2016) remind us that datasets will often contain surprises, even when they are ostensibly public and anonymous. Beyond the issue of joining data sets, there is the question of the ethics of experimentation (Crawford, 2014; Grimmelmann, 2015b). The most public example to date was the public furor over the Facebook "emotional contagion" study in 2014. After using large-scale A/B testing to manipulate the emotional valence of the news feeds of nearly 700,000 users, Facebook shared the results with then-Cornell social scientist Jeff Hancock, who co-published the study in the Proceedings of the National Academy of Science (Kramer et al., 2013). Susan Fiske, who edited the article for PNAS, relayed in public statements that the Cornell IRB had approved Hancock's role in the study because the dataset was 'pre-existing' as Facebook's data when he was first invited to participate in the analysis. His role in the study therefore did not technically rise to the standard of an 'intervention' in a human life that qualifies a study as human-subjects research requiring further review, and the Cornell IRB therefore granted approval (Meyer, R., 2014). It does not appear that Facebook used any independent review process to approve the research that created the 'pre-existing' dataset under question, nor would they be required to do so under the Common Rule as a private entity. Instead of quelling concern, this response ignited a broad debate about the ethics of such experiments (Auerbach, 2015; Crawford, 2014; Grimmelmann, 2015a; Meyer and Chabris, 2015; Waldman, 2014; Watts, 2014).

In an analysis of the Facebook emotional contagion controversy, Michelle Meyer argued that critics are mistaken if they examine only the antecedent (the "B") of A/B testing and not the precedent (the "A") (Meyer, 2015; see also Meyer and Chabris, 2015). In what she identifies as the "A/B illusion," we have a tendency to focus on the ethics of changes resulting from an experiment and not the prior state. The "A/B illusion" illustrates what is essentially

a variant of the naturalistic fallacy for the Big Data era: the way things *are* is the way that things *should* be and any change must be ethically interrogated. Meyer argues that it is equally important to ethically interrogate the precedent state as the antecedent state in order to avoid falling victim to this illusion. Yet she notes that the historically-situated codifications of research ethics are are calibrated to a different model of experimentation such that this logical parity remains largely invisible, and Big Data ethics debates tend to focus on the antecedent state alone. Thus, she links the future epistemic success of A/B testing, and its role in the Internet economy, to a regulatory environment that is not burdened by ill-fitting research ethics regulations that were historically designed to manage different scientific regimes.

While we largely agree that A/B testing is a loose fit for existing research ethics regulations, we argue that there are significantly different lessons to be drawn from those gaps. Meyer's argument hinges on the limitations of the practice/research distinction—and on this point we agree. As we have discussed above, much of research ethics regulation can be viewed as managing the line between physician-as-caregiver and physician-as-researcher. Yet, there is no clear analogue of the practice/research distinction in data science because the *practice* of data science is iterative *research*. Indeed, it is problematic to import wholesale the ethical standards regulating that distinction. But we do not see this as reason to reject the application of research ethics regulations (or other responsive methods and enforceable standards) in data science. Rather, it is possible to read research ethics through a different framing that emphasizes ethics regulations as a form of *community assent* that enables self-regulation. Although the particularities of current research regulations cannot be directly ported to data science, the history of biomedical research ethics regulation indicates that the success of the field will depend on collectively assenting to transparent, enforceable norms of trust, responsibility and accountability. We see formal research ethics regulation as a *route* to that goal rather than a goal in itself. It is crucial that research ethics norms are established with a nuanced and empirically-informed assessment of the potential harms of data—public, semi-public and private—and a critical understanding of emerging forms of human-subjects research.

Conclusion

Large-scale data experimentation in academia and industry is playing a significant role in shaping both scientific endeavor and much of everyday life. From social media platforms to city streets, data is being gathered and used to conduct experiments on the public. And yet there is very little research on how to identify, track and mitigate the risk imposed on people who are (often unwittingly) participants in these experiments. The current debate, and the HHS revisions as they are currently framed, might lock in potentially risky forms of research as exempt from review, and maintain a problematic sense that Big Data research does not directly impact people's lives.

Social scientists have a long and sometimes fraught relationship with the framing and reach of research ethics. As we have shown, this is due to the history of shaping ethics regulations around the epistemic conditions and particular scandals of biomedical research. Critical data studies should help articulate how new methods of knowledge production are co-constitutive with emergent ethical norms and modes of subjectivity (Jasanoff, 2004; Reardon, 2004; Thompson, 2013). Because the boundaries of human-subjects research are continually contested, it is crucial for new fields like data science to be attuned to the potential human impact of their work if they are to earn and maintain community trust. We argue that any move to exclude data science research from review, and more broadly, to consider it outside of human-subjects research, is thus premature and potentially dangerous. Rather, we propose that critical data studies contribute to a deeper understanding of data subjectivity, including an account of the fundamental responsibility that researchers have to care for the well-being of their subjects.

The changes proposed in the NPRM are claimed to be scaled toward empirical measurements of harm. But what is to be done with a field such as data science where practices for measuring and mitigating harms are still taking shape? What is "public" and "private" is not easily answerable by looking at the conditions of a database, but the proposed changes to the Common Rule appear to eliminate any formal point at which these questions could be asked. If adopted in a manner that does not allow for tracking the evolving risk profiles of data-intensive research, these new regulations could prematurely close off significant questions about data ethics. Both the NPRM and the National Academies report do recognize that risk profiles are rapidly changing with data-intensive research techniques, and suggest establishing an independent body capable of providing continuing advice to IRBs about how to measure and mitigate such risk (Committee on Revisions, 2014: 112–115). More accurate assessments of harms and risks are critical to ensure accurately and consistently assigning projects to the correct regulatory categories.

Finally, we should reject the belief that the risk borne by research subjects depends on what kind of data is obtained and how, rather than what is done with the data. In the context of data science, it simply

does not hold. Instead, large-scale data practices begin with the assumption that new insights—some extremely sensitive—can be generated through connecting previously disparate data sets. Thus, the Common Rule needs to reflect that even anonymous, public data sets can produce harms depending on how they are used. The best way to do this in academic settings remains the IRB. As for industry, there needs to be a more serious commitment to review and assessment of human data projects. Facebook, for example, responded to the public outcry about the emotional contagion experiment by setting up an internal review process for future experiments. Legal scholar Ryan Calo has argued that a body like the Federal Trade Commission could commission an interdisciplinary report on data ethics, and that those public principles could guide companies as they form small internal committees that review company practices (Calo, 2013). Polonensky et al. (2015) have similarly argued for a two-track ethics review model for use outside of the purview of the Common Rule that would blend internal and external perspectives. Dove et al. (2016) recently surveyed how research ethics committees have grappled with data-intensive research with "bottom-up" approaches when more traditional "top-down" approaches have fallen short. Others have also offered promising insights for integrating ethical reasoning into data science research and practice prior to the typical timing of formal ethical review (Shilton and Sayles, 2016; Steinmann et al., 2015; Tractenberg et al., 2015). We think these are valuable approaches going forward, with an emphasis on bringing data science practices into frameworks of trust and accountability. Rather than seeking to exempt entire classes of new and emerging research, we should be establishing more flexible and informed structures of review, both within the academy and in industry.

Acknowledgements

Declaration of conflicting interests

Funding

Notes

1. We will focus largely on the US context for the purpose of this paper. Although the specifics of the regulations differ in other nations, the practical and philosophical challenges posed by regulating Big Data with existing norms are similar elsewhere.
2. We are choosing not to repeat the name of the individual mentioned by the researchers, and the *Daily Mail*, given that it is personally identifying information which is unnecessary to the aims of this article. This is an ethical choice that we recommend applying more broadly in data science: only use the data that you need and that does not create risk for others.
3. In communication with the author, 3 March 2016.
4. In the USA, human-subjects research protections are governed by Title 45 Section 46 of the Code of Federal Regulations, known as "the Common Rule" because it applies commonly to research involving federally funded human subjects regardless of scientific discipline and across 15 separate federal signatory agencies. The current regulations were first published in 1981 by a handful of agencies and were revised to cover more agencies in 1991 in the format now known as the Common Rule.
5. According to Beauchamp, although the Belmont authors were focused primarily on recent scandals in medical research, they did not intend to imply that clinical practice should not be regulated or untested. He specifies the intensive collection of data about medical outcomes, and the best practices learning that can result from continual data analytics, as the type of activity that federal regulators and IRBs will find increasingly challenging to fit into the traditional framework (Beauchamp, 2011b).
6. Among them are: new uniform standards for informed consent forms, requirements for blanket consent for the secondary use of banked biospecimens, classifying bio-specimens as "human subjects" for the first time due to concerns about re-identification of samples, some accommodation of quality improvement studies and a lack of clarity on alternative IRB models that may be useful for many data scientists.
7. This analysis is echoed in the public comment on the NPRM published by the Council for Big Data, Ethics and Society in 2016 (see: Metcalf, 2016).
8. Regulatory agencies have provided some degree of clarification over time (Office for Human Research Protections, 2008a, 1995). But the National Academies' report on the Common Rule identifies lack of guidance from federal regulators to IRBs and researchers about how to empirically—rather than intuitively—measure the risks and benefits of SBE research as a source of "insurmountable barriers" for IRBs (Committee on Revisions, 2014: 66–67).
9. Importantly, the NPRM states that falling outside the regulatory scope of "human-subjects research" does not relieve researchers of ethical duties identified in the Belmont Report.

10. Previously, the Common Rule did not make a distinction between research data and data "collected for non-research purposes." This appears to be a change made to accommodate Big Data methods. The National Academies report made such a request so that researchers could have clear access to "found data" and "harvested data" located in public datasets. (Committee on Revisions, 2014: 43–47). In our view, the request to exclude all such research is questionable given the Committee's fairly extensive accounting of the dynamic and unpredictable risks posed by the rise of this type of research.

11. In clearer terms, the National Academies report on Common Rule revisions for the social sciences report makes the same recommendation: "New forms of large-scale data should be included as 'not human-subjects research' if all information is publicly available to anyone (including for purchase), if persons providing or producing the information have no reasonable belief that their private behaviors or interactions are revealed by the data, and if investigators using the data have no inter-action or intervention with individuals. Investigators must observe the ethical standards for handling such information that guide research in their fields and in the particular research context" (Committee on Revisions, 2014: 47).

References

Abbott L and Grady C (2011) A systematic review of the empirical literature evaluating IRBs: What we know and what we still need to learn. *Journal of Empirical Research on Human Research Ethics* 6(1): 3–19.

American Medical Association (2015) History of AMA ethics. American Medical Association: Our history. Available at: http://www.ama-assn.org/ama/pub/about-ama/our-history/history-ama-ethics.page (accessed 21 October 2015).

Andrejevic M (2014) Big data, big questions: The big data divide. *International Journal of Communication* 8(0): 17.

Annas GJ (1992) The changing landscape of human experi-mentation: Nuremberg, Helsinki and beyond. *Health Matrix: Journal of Law-Medicine* 2(2): 119.

Auerbach D (2015) The silicon tower. Slate. Available at: http://www.slate.com/articles/technology/bitwise/2015/05/facebook_study_why_silicon_valley_s_incursion_into_academic_research_is.html (accessed 6 November 2015).

Barocas S and Nissenbaum H (2014) Big data's end run around procedural privacy protections. *Communications of the ACM* 57(11): 31–33.

Bassett EH and O'Riordan K (2002) Ethics of internet research: Contesting the human subjects research model. *Ethics and Information Technology* 4(3): 233–247.

Beauchamp TL (2011a) Viewpoint: Why our conceptions of research and practice may not serve the best interest of patients and subjects. *Journal of Internal Medicine* 269(4): 383–387.

Beauchamp TL (2011b) The distinction between research and practice. Available at: https://www.youtube.com/watch?v=qPQ2HE2CfCA (accessed 21 October 2015).

Beauchamp TL and Saghai Y (2012) The historical founda-tions of the research-practice distinction in bioethics. *Theoretical Medicine and Bioethics* 33(1): 45–56.

Bowker GC (2005) *Memory Practices in the Sciences.* Cambridge, MA: MIT Press.

boyd d (2014) It's complicated: The social lives of networked teens. Yale University Press. Available at: http://dl.acm.org/citation.cfm?id=2584525 (accessed 6 November 2015).

boyd d and Crawford K (2012) Critical questions for big data. *Information, Communication & Society* 15(5): 662–679.

boyd d, Levy K and Marwick AE (2014) The networked nature of algorithmic discrimination. In: Gangadharan SP, Eubanks V and Barocas S (eds) *Data and Discrimination: Collected Essays, New America*, pp.43–57. Available at: http://www.newamerica.org/downloads/OTI-Data-an-Discrimination-FINAL-small.pdf (accessed 30 April 2016).

Brunton F and Nissenbaum H (2015) *Obfuscation.* Cambridge, MA: MIT Press. Available at: https://mitpress.mit.edu/books/obfuscation (accessed 27 October 2015).

Calo R (2013) Consumer subject review boards: A thought experiment. *Stanford Law Review Online* 66: 97.

Committee on Revisions to the Common Rule for the Protection of Human Subjects, Committee on National Statistics, et al. (2014) Proposed Revisions to the Common Rule for the Protection of Human Subjects in the Behavioral and Social Sciences. Available at: http://www.nap.edu/read/18614/chapter/1 (accessed 21 October 2015).

Crawford K (2014) The test we can—and should—run on Facebook. *The Atlantic.* Available at: http://www.theatlantic.com/technology/archive/2014/07/the-test-we-canand-shouldrun-on-facebook/373819/ (accessed 21 January 2015).

Crawford K and Schultz J (2014) Big data and due process: Toward a framework to redress predictive privacy harms. *Boston College Law Review* 55: 93.

Danyllo WA, Alisson VB, Alexandre ND, et al. (2013) Identifying relevant users and groups in the context of credit analysis based on data from Twitter. In: *2013 Third international conference on cloud and green comput-ing (CGC)*, Karlsruhe, 30 September–2 October 2013, pp.587–592. New York, NY: IEEE. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6686094 (accessed 24 February 2016).

Department of Health and Human Services (2009) Code of Federal Regulations Title 45 – Public Welfare, Part 46 – Protection of Human Subjects. 45 Code of Federal Regulations 46. Available at: http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html (accessed 19 October 2015).

Department of Health and Human Services (2015a) Notice of Proposed Rule Making: Federal Policy for the Protection of Human Subjects. Federal Register. Available at: http://www.gpo.gov/fdsys/pkg/FR-2015-09-08/pdf/2015-21756.pdf (accessed 21 October 2015).

Department of Health and Human Services (2015b) Notice of Proposed Rule Making: Federal Policy for the Protection of Human Subjects. Federal Register. Available at: https://s3.amazonaws.com/public-inspection.federalregister.gov/2015-21756.pdf (accessed 21 October 2015).

Dove ES, Townend D, Meslin EM, et al. (2016) Ethics review for international data-intensive research. Science 351(6280): 1399–1400.

Duster T, Matza D and Wellman D (1979) Field work and the protection of human subjects. The American Sociologist 14: 136–142.

Dwork C (2011) A firm foundation for private data analysis. Communications of the ACM 54(1): 86–95.

Dwork C and Mulligan DK (2013) It's not privacy, and it's not fair. Stanford Law Review Online 66: 35.

Franceschi-Bicchierai L (2015) Finding Muslim NYC Cabbies in trip data – Album on Imgur. Mashable. Available at: http://mashable.com/2015/01/28/redditor-muslim-cab-drivers/#0_uMsT8dnPqP (accessed 6 November 2015).

Frankel MS (1989) Professional codes: Why, how, and with what impact? Journal of Business Ethics 8(2–3): 109–115.

Fost N and Levine RJ (2007) The dysregulation of human subjects research. JAMA 298(18): 2196–2198.

Gaumnitz BR and Lere JC (2002) Contents of codes of ethics of professional business organizations in the United States. Journal of Business Ethics 35(1): 35–49.

Gawande A (2007) A lifesaving checklist. New York Times, 30 December. Available at: http://www.nytimes.com/2007/12/30/opinion/30gawande.html (accessed 30 April 2016).

Ghooi RB (2011) The Nuremberg Code – A critique. Perspectives in Clinical Research 2(2): 72–76.

Gitelman L (ed) (2013) Raw Data Is an Oxymoron. Cambridge, MA: MIT Press.

Grimmelmann J (2015a) The law and ethics of experiments on social media users. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. Available at: http://papers.ssrn.com/abstract=2604168 (accessed 19 October 2015).

Grimmelmann J (2015b) Ethical culture clashes in social media research. 2d. Laboratorium. Available at: http://2d.laboratorium.net/post/108480841510/ethical-culture-clashes-in-social-media-ressearch (accessed 6 November 2015).

Hauge MV, Stevenson MD, Rossmo DK, et al. (2016) Tagging Banksy: Using geographic profiling to investigate a modern art mystery. Journal of Spatial Science 61(6): 185–190.

Ioannidis JPA (2013) Informed consent, big data, and the oxymoron of research that is not research. The American Journal of Bioethics 13(4): 40–42.

Jackson SJ, Gillespie T and Payette S (2014) The policy knot: Re-integrating policy, practice and design in Cscw Studies of Social Computing. In: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, CSCW '14, New York, NY, USA: ACM, pp. 588–602.

Jasanoff S (2004) States of Knowledge: The Co-production of Science and the Social Order. London: Routledge.

Kaptein M and Wempe J (1998) Twelve Gordian knots when developing an organizational code of ethics. Journal of Business Ethics 17(8): 853–869.

Kass N, Pronovost PJ, Sugarman J, et al. (2008) Controversy and quality improvement: Lingering questions about ethics, oversight, and patient safety research. Joint Commission Journal on Quality and Patient Safety/Joint Commission Resources 34(6): 349–353.

Kitchin R (2014) Big data, new epistemologies and paradigm shifts. Big Data & Society 1(1): 2053951714528481.

Kosinski M, Stillwell D and Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences 110(15): 5802–5805.

Kramer ADI, Guillory JE and Hancock JT (2014) Experimental evidence of massive-scale emotional contagion through social networks. Proceedings of the National Academy of Sciences 111(24): 8788–8790.

Ledford H (2007) Human-subjects research: Trial and error. Nature 448(7153): 530–532.

Librett M and Perrone D (2010) Apples and oranges: Ethnography and the IRB. Qualitative Research 10(6): 729–747.

Mayer-Schönberger V and Cukier K (2013) Big Data: A Revolution That Will Transform How We Live, Work, and Think. New York: Houghton Mifflin Harcourt.

Meyer MN (2014) Misjudgements will drive social trials underground. Nature 511(7509): 265–265.

Meyer MN (2015) Two cheers for corporate experimentation: The A/B illusion and the virtues of data-driven innovation. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. Available at: http://papers.ssrn.com/abstract=2605132 (accessed 19 October 2015).

Meyer MN and Chabris CF (2015) Please, corporations, experiment on us. The New York Times, 19 June. Available at: http://www.nytimes.com/2015/06/21/opinion/sunday/please-corporations-experiment-on-us.html (accessed 19 October 2015).

Meyer R (2014) Everything we know about Facebook's secret mood manipulation experiment. The Atlantic. Available at: http://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/ (accessed 18 May 2016).

Metcalf J (2014) Ethics codes: History, context, and challenges. Council for Big Data, Ethics, and Society. Available at: http://bdes.datasociety.net/council-output/ethics-codes-history-context-and-challenges/ (accessed 21 October 2015).

Metcalf J (2016) Letter on proposed changes to the common rule. Council for Big Data, Ethics, and Society. Available at: http://bdes.datasociety.net/council-output/letter-on-proposed-changes-to-the-common-rule/ (accessed 11 January 2016).

Michael K and Miller KW (2013) Big data: New opportunities and new challenges [Guest editors' introduction]. Computer 46(6): 22–24.

National Commission for the Protection of Human Subjects, of Biomedical and Behavioral Research and The National Commission for the Protection of Human Subjects (1979) The Belmont Report: Ethical Principles and Guidelines

for the Protection of Human Subjects of Research. Available at: http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html (accessed 21 October 2015).

Neuhaus F and Webmoor T (2012) Agile ethics for massified research and visualization. *Information, Communication & Society* 15(1): 43–65.

Nissenbaum H (2009) *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford University Press.

Nissenbaum H (2011) A contextual approach to privacy online. *Daedalus* 140(4): 32–48.

Nuremberg Code (1949) Trials of war criminals before the Nuremberg Military Tribunals under Control Council Law No. 10. U.S. Government Printing Office. Available at: http://www.hhs.gov/ohrp/archive/nurcode.html (accessed 21 October 2015).

Office for Human Research Protections (1995) Exempt research and research that may undergo expedited review. Available at: http://www.hhs.gov/ohrp/policy/hsdc95-02.html (accessed 27 October 2015).

Office for Human Research Protections (2008a) Secretary's Advisory Committee on Human Research Protections Letter to HHS Secretary. Available at: http://www.hhs.gov/ohrp/sachrp/sachrpletter091808.html (accessed 27 October 2015).

Office for Human Research Protections (2008b) OHRP Statement Regarding The New York Times Op-Ed Entitled 'A Lifesaving Checklist'. Health and Human Services. Available at: http://archive.hhs.gov/ohrp/news/recentnews.html#20080115 (accessed 21 October 2015).

Polonetsky J, Tene O and Jerome J (2015) Beyond the common rule: Ethical structures for data research in non-academic settings. *Colorado Technology Law Journal* 13: 101.

Pronovost P, Needham D, Berenholtz S, et al. (2006) An intervention to decrease catheter-related bloodstream infections in the ICU. *New England Journal of Medicine* 355(26): 2725–2732.

Reardon J (2004) *Race to the Finish: Identity and Governance in an Age of Genomics*. Princeton, NJ: Princeton University Press.

Rhodes R, Azzouni J, Baumrin SB, et al. (2011) De minimis risk: A proposal for a new category of research risk. *The American Journal of Bioethics* 11(11): 1–7.

Shea C (2000) Don't talk to the humans: The crackdown on social science research. *Lingua Franca* 10(6): 1–6.

Shilton K and Sayles S (2016) 'We Aren't All Going to Be on the Same Page About Ethics': Ethical practices and challenges in research on digital and social media. In:

Proceedings of the 49th Hawaii international conference on system sciences, Kauai, HI. Available at: https://terpconnect.umd.edu/~kshilton/pdf/ShiltonSaylesHICSSpreprint.pdf (accessed 9 February 2016).

Silberman G and Kahn KL (2011) Burdens on research imposed by Institutional Review Boards: The state of the evidence and its implications for regulatory reform. *The Milbank Quarterly* 89(4): 599–627.

Steinmann M, Shuster J, Collmann J, et al. (2015) Embedding privacy and ethical values in big data technology. In: *Transparency in Social Media*. Switzerland: Springer, pp.277–301. Available at: http://link.springer.com/chapter/10.1007/978-3-319-18552-1_15 (accessed 24 February 2016).

Thompson C (2013) *Good Science: The Ethical Choreography of Stem Cell Research*. Cambridge, MA: The MIT Press.

Thompson DA, Kass N, Holzmueller C, et al. (2012) Variation in local Institutional Review Board evaluations of a multicenter patient safety study. *Journal for Healthcare Quality* 34(4): 33–39.

Tockar A (2015) Riding with the Stars: Passenger Privacy in the NYC Taxicab Dataset. Neustar Research. Available at: http://research.neustar.biz/2014/09/15/riding-with-the-stars-passenger-privacy-in-the-nyc-taxicab-dataset/ (accessed 6 November 2015).

Tractenberg RE, Russell AJ, Morgan GJ, et al. (2015) Using ethical reasoning to amplify the reach and resonance of professional codes of conduct in training big data scientists. *Science and Engineering Ethics* 21(6): 1485–1507.

Waldman K (2014) Facebook's unethical experiment. *Slate*. Available at: http://www.slate.com/articles/health_and_science/science/2014/06/facebook_unethical_experiment_it_made_news_feeds_happier_or_sadder_to_manipulate.html (accessed 6 November 2015).

Walther JB (2002) Research ethics in internet-enabled research: Human subjects issues and methodological myopia. *Ethics and Information Technology* 4(3): 205–216.

Watts DJ (2014) Lessons learned from the Facebook study. *The Chronicle of Higher Education Blogs: The Conversation*. Available at: http://chronicle.com/blogs/conversation/2014/07/09/lessons-learned-from-the-facebook-study/ (accessed 17 May 2016).

World Medical Association (1964) World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. World Medical Association. Available at: http://www.wma.net/en/30publications/10policies/b3/17c.pdf (accessed 21 October 2015).

Zwitter A (2014) Big data ethics. *Big Data & Society* 1(2): 2053951714559253.

This article is a part of Special theme on Critical Data Studies. To see a full list of all articles in this special theme, please click here: http://bds.sagepub.com/content/critical-data-studies.