

Internet Users as Vulnerable and at-Risk Human Subjects: Reviewing Research Ethics Law for Technical Internet Research

Thesis submitted in partial fulfilment of the requirements for the degree of DPhil in
Information, Communication and the Social Sciences
at the Oxford Internet Institute, University of Oxford



Bendert Zevenbergen
Oxford Internet Institute, Green Templeton College

Supervisors:
Dr. Victoria Nash, Oxford Internet Institute, University of Oxford
Dr. Joss Wright, Oxford Internet Institute, University of Oxford

Word count: 93,531

Acknowledgements

I owe tremendous gratitude to a huge amount of people. Primarily, I want to thank everyone at the Oxford Internet Institute, a wonderful group of scholars, students, and administrative staff who were among the first to pioneer multidisciplinary sociotechnical research. Specifically, I want to thank my supervisors Vicki Nash, Joss Wright, and Ian Brown. Their critical and constructive approaches to my work allowed me to write a thesis that I would not have been able to produce otherwise. It's these interactions that make a D.Phil research so valuable and force you to grow as a person, in terms of intellect and resilience. The same amount of gratitude goes to Green Templeton College, where I spent many years working, meeting new friends, engaging in sports, and enjoying the best food Oxford University has to offer. Principal Denise Lievesley and her partner Roland Rosner have made the experience especially memorable. I will not thank all my friends here individually, as it would cover several pages. The thesis would not have been possible without the generous support of the Open Technology Fund and Google Research. While they did not sponsor the thesis directly, the side projects they enabled me to work on, inspired me to change the focus of my thesis to the problems with research ethics evaluation of technical Internet research. I also spent three years at Princeton University, working on other projects while writing this thesis. My experience working with the Center for Information Technology Policy and the Center for Human Values were invaluable. I would like to acknowledge in particular Melissa Lane, Ed Felten, and Nick Feamster for the tremendous experience in New Jersey. Also, the support staff for academic software such as DevonThink, Scrivener, and Bookends have saved this work more than once. Last, but not least, I want to thank my parents, Nicole and Christiaan Zevenbergen. They provided me with the education and free personality to even consider spending several years working on a thesis. I'm also deeply thankful for the six months in 2019 and 2020 I spent at their home in Arnhem to finally finish this thesis once and for all.

Table of contents

(1) Introduction.....	7
Research overview	7
Relevance of research topic	9
Personal experience and motivation	12
Relevant literature.....	15
Research ethics	15
Law, policy, and socio-legal studies.....	16
Ethics.....	18
Position within literature	19
Methodology.....	22
Layout of the thesis.....	31
(2) Historical Emergence of Research Ethics.....	35
Introduction	35
Part 1: Researchers' Roles, Relationships, and Responsibilities.....	37
Part 2: Models of Research Ethics.....	42
Atrocities	43
Reactionary Regulation: The Human Subject Model.....	44
Codification into Law	46
The United States	47
European Union	48
Overview of Formal Regulatory Structure.....	53
Other Models of Research Ethics	55
Part 3: Concepts of the Common Rule.....	57
Research.....	58
Human Subjects.....	60
Risk of Harm vs. Benefits.....	62
'No More Than Minimal Risk'	64
Assessment.....	65
Autonomy	66
Autonomy in Society	67
Relevance for Research Ethics.....	68
Informed Consent	69
Vulnerable Persons	73
Deception	76

Privacy	78
Existing and Publicly Available Data	81
Confidentiality	81
Anonymity	82
Part 4: Law, Ethics, and Research Themes	83
Law	84
Traditional role of law	84
Expressive and communicative function of law	86
Ethics Frameworks.....	87
Ethics of care: Background	94
Ethics of Care: Research Themes	96
(3) Conceptual Development of Internet Research Ethics	107
Introduction	107
Ethics review in a new environment.....	109
The Internet as a research domain.....	111
Challenge to information laws.....	114
Social Science Internet Research Ethics.....	117
Controversies	119
Challenges to existing research ethics concepts	123
Human subjects	124
Minimal risk.....	126
Autonomy	128
Privacy	133
Existing social science Internet research ethics guidelines	140
Technical Internet research ethics.....	143
Research ethics guidelines	147
Challenges to the research ethics framework	150
Human subjects	150
Minimal risk.....	152
Autonomy	154
Privacy	155
Towards a New Internet Research Ethics?	161
Conceptual framework (1).....	165
(4) Case Study 1: Password Dumps in Research	170
Introduction	170

Contextualising the case study paper	171
Research on illicit data	171
Password 'dumps'	174
Aims and Purpose of password research	175
Description of the paper.....	184
Common Rule analysis.....	187
Research and approval.....	189
Human Subjects.....	191
Exemptions from review.....	193
Waiver for informed consent.....	194
Structured case research cycle analysis	197
Are password dumps “existing” and “publicly available” data?.....	198
Waiver for informed consent requirement.....	201
Conceptualising harm	207
Conceptual Framework: 2nd Iteration.....	213
(5) Case Study 2: Censorship measurement.....	221
Introduction	221
Contextualizing the case study paper.....	222
Internet measurement	223
History and objective.....	223
Censorship measurement	228
Description of paper	233
Common Rule Analysis	242
IRB review	243
Program committee	246
Alternative IRB process.....	248
Structured case research cycle analysis	259
Research.....	260
Human subjects	262
Development of the human subject concept.....	262
Common Rule requirements	264
Distance and the disintermediated research relationship.....	268
Informed consent	269
Autonomy	272
No more than minimal risk of harm	275

Actors and attackers	277
Intrinsic harms.....	279
Extrinsic harms	282
Conceptual Framework: 3rd Iteration	290
(6) Analysis	299
Literature-based scrutiny	299
Human Subjects.....	308
Common Rule.....	308
Challenges from conceptual framework	310
Articulation of a new provision.....	314
Autonomy and vulnerability.....	317
Common Rule.....	318
Challenges from conceptual framework	322
Articulation of a new provision.....	334
Harm and privacy	337
Common Rule.....	337
Challenges from conceptual framework	340
Articulation of a new provision.....	349
(7) Conclusion.....	351
Answers to research questions.....	354
Practical contributions.....	357
Research Ethics Committees (and Institutional Review Boards).....	357
Technical Internet researchers.....	358
Policymakers.....	358
Ethics of care scholarship.....	359
Future research.....	359
References.....	362

(1) Introduction

Research overview

In June 2014, various media outlets revealed that the popular social networking website Facebook had manipulated Internet users' emotions for their own scientific ends.¹ The CEO of another online service—a dating website called OkCupid—responded to the public outrage following the revelation, stating that:

“We noticed recently that people didn’t like it when Facebook ‘experimented’ with their news feed. Even the FTC is getting involved. But guess what, everybody: if you use the Internet, you’re the subject of hundreds of experiments at any given time, on every site. That’s how websites work”
(Rudder, 2014).²

This quote indicates that it is common practice for companies and other powerful actors on the Internet to experiment with unsuspecting Internet users without their explicit knowledge. The objectives of this type of research can range from scientific experimentation and product optimisation to simply testing the limits of what is possible on the Internet. Academics, too, participate in these practices by collecting and reusing data collected on the Internet in large quantities. However, most of the

¹ See, for example: Adrienne LaFrance (2014), “Everything We Know About Facebook's Secret Mood Manipulation Experiment”, *The Atlantic*, available at: <https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/> and Reed Albergotti and Elizabeth Dwoskin (2014), “Facebook Study Sparks Soul-Searching and Ethical Questions”, *The Wall Street Journal*, available at: <https://www.wsj.com/articles/facebook-study-sparks-ethical-questions-1404172292>

² The original blogpost has since been deleted, though it is reported on in various sources, such as in the Guardian’s article title “OKCupid: we experiment on users. Everyone does”, available at: <https://www.theguardian.com/technology/2014/jul/29/okcupid-experiment-human-beings-dating>.

Internet's unwitting research subjects are not typically aware of how the technical aspects of the Internet function, which data trails they generate as a user, how this data may be collected, and that their devices or online experience may be subject to experimentation.

The aim of this thesis is to assess the extent to which the research ethics frameworks imposed on technical Internet researchers in academic institutions—in disciplines such as computer science and network engineering—are fit-for-purpose, whether they yield morally satisfactory outcomes, and, if not, how they should be reinterpreted or reviewed. Typically, research ethics review boards apply the human subjects model of research ethics to most academic research, which is a model that was originally developed for experimentation in the biomedical sciences to protect the human research subject. It may be intuitively appealing to apply this existing human subjects model to technical research, too, since it also served its purpose in other disciplinary domains to some extent. However, this thesis demonstrates that transposing a research ethics framework onto technical Internet research may be problematic, because the assumptions underlying the concepts of the law are challenged by the inherent, novel, and unique characteristics of the Internet. This thesis therefore serves as a thought-experiment on the intersection of law, ethics and technology.

The main, overarching research question of this thesis is, therefore:

1. How ought technical Internet research with unsuspecting data subjects be assessed from the perspective of research ethics?

The main research question is supported by two secondary research questions, which are answered in the case study chapters of this thesis:

2. Which concepts and principles of research ethics frameworks are challenged specifically by technical Internet research with unsuspecting data subjects?
3. How might these concepts be reinterpreted for the purposes of research ethics evaluation of technical internet research when considered through the appropriate ethical framework?

The thesis provides a much-needed review of research ethics procedures in computer science. It makes four distinct contributions:

1. Insight for Institutional Review Boards (IRBs) and other research ethics gatekeepers to aid the interpretation of the current research ethics frameworks when reviewing technical Internet research studies,
2. Guidance for scholars designing and conducting technical Internet research when assessing the research ethics requirements of their studies,
3. A socio-legal roadmap for legislative or self-regulatory review by policymakers or other persons responsible for research ethics frameworks for technical Internet research,
4. A novel assessment in the developing philosophical literature on the ethics of care.

Relevance of research topic

Research ethics frameworks are one of the cornerstones of academic freedom, which in turn is the key concept that defines the professional identity of academic scholars.

A sufficiently strong and appropriate research ethics framework breeds trust in the academic profession and allows academics to operate with a relative amount of freedom. Academics rely on the trust of policymakers, and by extension that of the societies following those policies, to conduct their work. They are financially dependent to a large extent on public funds, but they also exert significant power in society. For example, academic work typically underpins policy debates and stimulates progress and innovation in the economy. At the same time, as academics possess superior knowledge in specific domains compared to the general population, society needs to trust that academics do not use that power to put their research subjects at risk of harm. In the absence of clear rules, a power imbalance may emerge between researchers and the data subjects upon whose trust the research profession depends. By imposing research ethics frameworks on researchers, society gains some degree of trust that the research profession can police themselves and that people will not be subjected or fall prey to the whims of researchers.

Research ethics frameworks have been developed worldwide, typically as a policy response to atrocities in the name of scientific discovery. In the past, researchers have purposefully spread diseases, subjected people to unimaginable physical harm or torturous mind games, and killed people to acquire new knowledge. The most well-known atrocities are in the biomedical and behavioural sciences, such as those conducted in Nazi concentration camps. As a result of these atrocities, international legal instruments such as the Nuremberg Code and the Declaration of Helsinki now underpin the development of national and discipline-specific research ethics frameworks. The phrasing of these documents and their resulting frameworks

reflects the thought process and assumptions that their drafters had about potential atrocities in research and how to curtail them.

The widespread adoption of the Internet has led researchers to use the global digital network as a venue, tool, or subject of research. The Internet is particularly relevant for social science research because of its social significance in (more or less) all domains of the modern human experience. As a technical phenomenon, the Internet has been the subject of ongoing technical studies and experimentation for several decades. To some extent, scholars and ethics boards have struggled to apply the existing research ethics frameworks to the new types of research enabled by the Internet. Moreover, when technical Internet research enters the social domain (e.g., by experimenting on live data points that are, in fact, human Internet users), ethics boards have typically struggled to fit the human element into the existing frameworks. The result is that studies with ethically dubious outcomes or impacts have not been reviewed, certain concepts have been interpreted poorly, or beneficial research has been halted for dubious reasons.

Many examples are discussed throughout this thesis. To give a sense of the larger picture, researchers at Facebook and Cornell University spread positive and negative emotions on Facebook, anonymous researchers infiltrated around half a million devices (including security door and industrial control systems) just to see what they could do by imitating malicious code, and researchers at Stanford collected a vast amount of digital portrait pictures from specific dating apps to train their neural network to recognise sexual preferences in human faces. These particular examples are not case studies in the thesis, but are exemplary of the questionable practices Internet researchers engage in. Further, almost every computer scientist has a story

about how their research proposal was rejected by the ethics board simply because the parties could not agree on the ethical considerations of the study. Whether such complaints are justified is not always clear, although it does indicate a loss in disciplinary translation of the required or expected ethical standards. This thesis assesses the loss in translation when the research ethics frameworks of the biomedical and behavioural sciences are transposed to Internet research.

Personal experience and motivation

My interest in this topic is the result of a serendipitous journey as a young scholar. I returned to academia after a few years of practising law and working in policy. My work in law and policy revealed to me that there is a significant knowledge gap with regards to social, philosophical, and legal concepts between the engineers who create the technologies that support Internet-enabled and mediated experiences, and the policymakers who regulate society. I decided to pursue my Ph.D (or D.Phil), to work across disciplines and learn their various languages and approaches. My end goal is to study how a lawyer can effectively work with engineers to develop technology in a way that respects age-old principles of society, rather than myopically optimising for a limited range of values, such as efficiency, speed, or scalability of individual systems, that are typically incentivised in engineering disciplines.

Soon after starting the first iteration of my doctorate, I was invited to work with several academic computer science communities. Certain technical Internet research communities and providers of research infrastructure struggled to communicate the ethical appropriateness and limitations of particular studies in their fields. The various research communities involved in Internet research voiced concerns about the difficulty of upholding ethical standards with the tools provided by legislators or

by discipline-specific guidelines. Programme committees at computer science conferences repeatedly asked me to review controversial papers through the lens of ethics, speak at their conferences when major ethical issues arose in their communities, and develop new guidelines that would be more applicable to their work than those provided by legislators.

The most notable results of this side-project are a co-authored set of specific guidelines for technical Internet measurements in which data is collected from mobile phones (Zevenbergen, Brown, Wright, & Erdos, 2013) and a broader set of guidelines to direct conversations between computer scientists and their ethical gatekeepers (Zevenbergen, 2016). Several computer science conferences referred authors and programme committees to these guidelines to guide their decision-making on issues concerning research ethics. Another aim of this project was to build a network of experts in academia, government, NGOs, and business who come from diverging theoretical and practical disciplines but share an interest in computer science research ethics. The work of this side-project became the main focus of my doctorate after its completion. The following three revelations led me to change my thesis topic to research ethics in computer science.

First, the Internet's unique characteristics give rise to new social phenomena and contravene commonly held assumptions about information. Large amounts of data revealing intimate details of distant people's lives can be easily collected on the Internet, and the environments of those distant people can be altered by the technically literate. The very nature of information has changed; therefore, the various pieces of legislation that regulate the uses of information need to be reviewed. Research ethics frameworks are among the information laws that should

be reconsidered in light of sociotechnical developments. To underpin this legislative review, it is useful to conduct a socio-legal study of the underlying technical shift and its impact on Internet users, researchers, and the application of the law.

Second, the disciplines of computer science and network engineering chiefly study technical artefacts. From a scientific point of view, these disciplines are not primarily interested in the human users of the systems on which they experiment. However, because the Internet is widely used to mediate many aspects of human life, technical studies will most likely implicate human Internet users in some way. Such researchers are not accustomed to assessing their work through a human subject lens and may be unaware of the fact that they could impact a large amount of people. One of the most notable anecdotes from the workshops organised prior to this thesis concerns a person who was instrumental in the early development of the Internet,³ who said:

"Well, we invented the Internet precisely so we wouldn't need to speak to anyone!"

As a result of recent controversies, however, various computer science and network engineering research communities have started to think about what it means to collect data from devices that belong to individual, identifiable persons. This is a noteworthy paradigm shift in computer science.

Third, as a cornerstone of academic freedom and social trust in the sciences, research ethics standards in computer science and other technical Internet research are crucial

³ The convenings were governed by Chatham House rules, so the person will remain nameless. Several people matching this description were present.

for a sustainable and socially legitimate development of the Internet. Billions of Internet users appear to be subjected to experimentation on a daily basis, by technically savvy researchers who may not have the interests of these users in mind. It is therefore particularly in the disciplines of computer science and network engineering where a well-informed and evidence-based review of the applicable research ethics framework is required, to uphold researchers' ethical responsibilities in this new research domain.

Relevant literature

This research project grew out of a multi- and interdisciplinary project on determining suitable research ethics approaches in computer science. This thesis therefore, necessarily, builds on several disciplinary literatures, such as law (including socio-legal studies), ethics, and computer science, as well as the literature that exists on their intersections. For example, research ethics frameworks are typically mandated by or even enshrined into law, or they are imposed as a softer regulatory instrument by academic conferences, universities, and their separate departments. The field of research ethics itself, however, is underpinned by ethical theories such as consequentialism and deontology or interpreted using a variety of other frameworks. The methodology of and approach to this study are influenced by socio-legal studies, which are also known as empirical legal studies. This section gives a brief, introductory, and high-level overview of the main traditions of scholarship within which this study is situated. The next two chapters discuss the literature of the various disciplines in greater depth.

Research ethics

The literature on research ethics exists in several formats and disciplines. For example, legal scholars such as James Grimmelmann assess particular studies

through the lens of applicable research ethics law and discuss whether the outcomes are appropriate or whether the frameworks need review (Grimmelmann, 2015).

Others collect and analyse data about the research ethics review procedures;

Elizabeth Buchanan and Charles Ess, for example, examined the technical literacy of ethics review boards at American universities (Buchanan, E. A. & Ess, 2009).

Practical ethicists such as Solomon Benatar and Peter Singer take aim at specific concepts within research ethics and set down the moral foundations for their own interpretation (Benatar & Singer, 2000). Scholars embedded in particular disciplines have also reflected on their experiences of applying research ethics principles to their work, such as Mark Allman and Craig Partridge's discussion of harm in technical Internet research (Partridge, C. & Allman, 2016). The wide range of disciplines involved in studying and commenting on research ethics is reflected in the journal articles published in various academic journals on the subject.⁴

Law, policy, and socio-legal studies

This thesis takes the US Common Rule as a legal subject of analysis. This law is enacted in the U.S. Department of Health and Human Services' Title 45 CFR 46 (Public Welfare) Subparts A, B, C and D.⁵ Contrary to most other jurisdictions, the US enshrined its research ethics framework and review procedure into federal law. It was developed as a response to the development of international treaties and certain infamous atrocities in the name of research, such as the Stanford prison experiment, the CIA's Project MKUltra, and Willowbrook State School. In these cases, unwitting research subjects were subjected to a very high risk of substantial harm (as is

⁴ See for example *Research Ethics* <<https://journals.sagepub.com/home/rea>>, *Journal of Empirical Research on Human Research Ethics* <<https://journals.sagepub.com/home/jre>>, *Ethics & Human Research* <<https://onlinelibrary.wiley.com/journal/25782363>>, and *Journal of Academic Ethics* <<https://link.springer.com/journal/10805>>.

⁵ Available at: <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html>.

discussed in the second chapter of this thesis). This law resembles research ethics frameworks imposed on researchers in other countries, in part because it is based on the Declaration of Helsinki. It is therefore a useful case in point for a socio-legal study.

The Common Rule has guided research in the biomedical, behavioural and later the social sciences relatively well, avoiding major ethical lapses, inappropriate experimentation, and the types of atrocities that led to the enactment of the law. Of course, scholars debated aspects of the Common Rule, such as Connie Varnhagen questioning the usefulness of informed consent (Varnhagen, et al., 2005) and Rosamund Rhodes challenging the assumption of data subjects' autonomy being the underlying aim of research ethics law (Rhodes, 2005). To some extent, however, it can be argued that scholars have successfully adapted the Common Rule's interpretations to fit the moral needs of novel research areas subjected to the law's concepts and review procedure.

This legislative success story may have come to an end with the need for a review of technical Internet research that involves human subjects. The ethical lapses of Internet researchers have been documented in newspaper headlines. Interestingly, studies that were considered controversial in the press were typically either found to be exempt from a research ethics review or went through the review procedure and were allowed to proceed.

This raises the concern that existing law may not be able to contain the behaviour that it is supposed to regulate. In fact, philosopher James Moor states that new developments in computing technology often leave a "policy vacuum" regarding the

appropriate use of computers (and their networks), since either there are no policies on the subject or the existing policies are inadequate (Moor, 1985). He refers to the resulting situation as a "conceptual muddle" between technical reality and the law in the books. He suggests that the emergence of a policy vacuum is typically the result of a conceptual vacuum, or a disparity between new technical capabilities and the terms and definitions used to regulate them. Moor calls for the development of conceptual frameworks to aid in the analysis of these technologies, in order to formulate appropriate regulatory policies. This thesis takes up the challenge of identifying the extent to which technical Internet research is currently conducted within a policy vacuum and how the conceptual framework for analysis should be constructed.

Ethics

This thesis is not primarily concerned with philosophy or ethics. However, research ethics frameworks are underpinned by ethical values, principles, their tensions, and their trade-offs. Hence the name research *ethics*. Similarly, laws, legal frameworks, and policies are typically also guided to a certain degree by ethics. This is also true for the Common Rule.

Many ethical frameworks and reasoning methodologies exist, but certain dominant ideas are especially relevant for this thesis. Abstract and universal approaches such as consequentialism (or utilitarianism) and deontology consider humans to be independent and rational actors who seek a fair and equitable society.⁶ These frameworks are useful for crafting legislation or policy and are recognised by scholars in the various research ethics laws and policies that exist today. Contextual

⁶ Ethics frameworks are defined and explored in more depth in the next chapter.

and relational ethical frameworks as offered by virtue ethics and the ethics of care are more useful for assessing the morality of actions within concrete, real-life contexts; these frameworks thus guide individual behaviour and factor in emotions such as empathy, sensitivity, wisdom, and courage. This thesis uses specifically the ethics of care as an analytical lens for the interpretation of concepts, because its focus on relationship and power dynamics is well-suited for an analysis of the responsibility of academics towards their research subjects. The theory and its component parts that are relevant for this thesis are described in the fourth part of Chapter 3.

Position within literature

Law

Legal analyses of Internet research ethics characteristically stay within the boundaries of existing law. The insights from legal analyses are useful for identifying the weak or ambiguous aspects of a legal framework, but this thesis aims to go beyond the confines of current law. A socio-legal approach is more suitable for this study because the discipline employs empirical methodologies to test the function, legitimacy, and suitability of an existing law in the real world. Laura Beth Nielsen captures the contribution of socio-legal studies when she states,

“To be effective, the law must be empirically examined in the real world and insights gleaned must inform law-makers through some sort of feedback mechanism.” (Nielsen, 2010, p. 23)

This study creates the feedback loop described by Nielsen by developing, testing, and strengthening a conceptual framework for technical Internet research ethics, as suggested by Moor.

Internet research ethics

Scholars in the social sciences have conducted research with human participants on or via the Internet for more than two decades. Due to their familiarity with research ethics procedures in their work, scholars developed a sizable corpus of literature and debate on the issues that arise when experimenting on the Internet. This work is well summarised in the Association of Internet Researchers' Ethics Guidelines.⁷ The lessons learned by social scientists are only partly directly applicable to the issues faced by technical Internet researchers, however, though they do form a fruitful starting point for this thesis. Their ideas and lessons form the first iteration of the conceptual framework developed in this thesis.

The technical Internet research community has fewer but still notable literature and debates on the research ethics issues that emerge in their work. Several authors have contributed thoughtful analyses of the ethical issues that arise when conducting experiments on the Internet, although they provide little guidance for researchers on the ground. In 2012, several scholars developed the Menlo Report for the US Department of Homeland Security. The aim of this document was to reinterpret the Belmont Report—a foundational report that paved the way for enactment of the US Common Rule—for information and communication research. While it is a thought-provoking intellectual exercise, technical researchers did not find the Menlo Report to be directly helpful as a guide for navigating the ethical tensions they face. To

⁷ Available at <<https://aoir.org/ethics/>>.

illustrate the lack of guidance on ethical matters for technical researchers and their ethical gatekeepers, a computer scientist once confided the following in me:

"Honestly, we often just throw in as much technical jargon to confuse the IRB, so we don't need to go through their review processes."

Ethics of care

The ethics of care is a normative theory that arose out of criticism by a group of (feminist) scholars about the existing and purely rational, abstract, and universalist ethical frameworks that were commonly applied in academia and research (Held, 2005). Moral theories such as consequentialism and deontology take an abstract and impersonal view of situations and consider actions without regard for the actor's personal context. The ethics of care brings the relationships, dependencies, and power structures experienced by the actor and those who they are responsible for to the foreground. The theory began to develop with an appreciation of the importance of private caring relationships for ethical reasons. It has since expanded to a large variety of private and public contexts, including international relations, the welfare state, and political theory (Bubeck, 1995). The ethics of care was initially developed in the 1970s and is still a work in progress. While the theory takes an alternative view of morality from the more common and masculine dominated theories, it is not synonymous with feminist ethics (Sander-Staudt, 2011). This section briefly describes its area of concern, primary values, and relevance for a thesis on research ethics, and will be elaborated upon in the next chapter.

The ethics of care focuses on the moral significance of relationships. More specifically, the theory assesses caring relationships in contexts where power dynamics are in play and an actor has been assigned the role—willingly or not—of

caregiving for more vulnerable and dependent persons. The ethics of care challenges oppressive social arrangements and imposes duties of care of other, more vulnerable people onto those in power. As a result, this theory prioritises values such as empathy, sensitivity, sympathy, responsiveness, attentiveness, and the resourceful interpretation of another person's needs. In contrast, a negative relationship according to the ethics of care would be characterised by the powerful actor's domination, their exploitation of the more vulnerable person, and mutual mistrust and hostility.

Academic researchers are typically situated within a web of relationships that includes their research participants, their colleagues, and their students. Further, the academic freedom they enjoy is predicated on their relationship with a society in which policymakers and citizens trust the scientific community to provide useful knowledge without using their power (in terms of knowledge and freedom) to cause harm. Therefore, the ethics of care offers a particularly useful framework for assessing the responsibilities of scientists and interpreting whether or not the existing obligations of research ethics promote good relationships.

Methodology

This thesis seeks to evaluate the usefulness and appropriateness of the legally mandated US human subject framework for research ethics review (specifically the Common Rule) and its component concepts in light of technical Internet research. To inform IRBs, researchers, and policymakers about how research ethics obligations for technical Internet researchers ought to be assessed, the deficiencies of the current framework must be laid bare. I use a combination of methodologies to connect new

research ethics issues with existing approaches and answer the research questions systematically.

This section is the only methodology section in this thesis. It is relatively short, because qualitative socio-legal methods are typically not as complex as quantitative approaches in social science. This subsection first clarifies the socio-legal traditions upon which this study builds. The second subsection describes the structured case methodology, which is used to develop the conceptual framework suggested by James Moor for research ethics law in the new context of computer science. This subsection also discusses the case selection criteria and the method of analysis during the case studies. The last subsection describes the final iteration of the structured case methodology, which is a literature-based examination of the resulting conceptual framework.

Socio-legal studies

Socio-legal studies, similar to disciplines such as law and society or empirical legal studies, focus on what the law does, rather than what it is, its nature, and its role in the institutions that govern society (Silbey & Sarat 1987; Cotterrell, 1997). This discipline assesses the difference between law in books and law in action (Pound, 1910) and attempts to understand the reasons that certain laws are abided by, while others are not (Tyler, 2006). Socio-legal studies therefore differ distinctly from traditional, black letter approaches to law, as socio-legal approaches offer an external point of view as opposed to the internal view of academic and practicing lawyers (Van Velthoven, 2016).

The uneven pace at which society, technology, science, and law advance (Ogburn, 1922; Marchant, Allenby, & Herkert, 2011) affords socio-legal studies a particularly useful role in informing future law-making, ensuring that law and science develop in tandem both appropriately and meaningfully (Jasanoff, 2007). Using a socio-legal approach rather than a black letter approach allows me to interrogate the apparent policy vacuum in research ethics created by Internet research.

Structured case methodology

A case study methodology is particularly suited to address the research questions put forward in this thesis because it examines specific cases to illuminate why decisions were made, how they were implemented, and with what result, without attempting to control the events of the case at hand (Schramm, 1971; Yin, 2017). Case studies have been considered to be especially useful for fields characterised by rapid changes in practice, such as technical Internet research (Markus, 1997).

This thesis employs a *structured case* methodology, which is a variation on the traditional case study methodology that was initially developed to test and enhance theories by researchers studying information systems (Carroll & Swatman, 2000).

This methodology constructs an analytical model, including a conceptual framework, several case study iterations, and a literature-based critique of the findings. It is the replication within a multiple case-study approach that provides the richness of examples and analyses which are necessary to drive theory-building from an initial conceptual framework (Eisenhardt & Graebner, 2007). For the purpose of this thesis, however, the focus on building theory is replaced by a focus on evaluating a legal framework and proposing new legal provisions. This methodology is well-suited to the thesis' broader aim of creating a feedback

mechanism for policymakers reviewing research ethics frameworks for technical Internet research because it assesses the status quo through the lens of real cases. The iterative case study approach of the structured case methodology is illustrated below as a “spiral of understanding”:

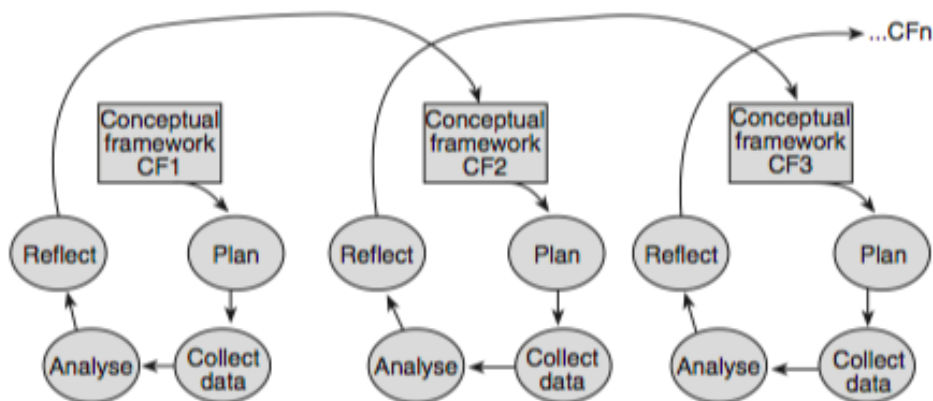


Figure 2 A spiral towards understanding (Carroll & Swatman 2000, p. 6)

This section briefly describes where each part of this spiral is located in this thesis. First, the conceptual framework (CF1) comprises our current understanding of research ethics as legislated in the Common Rule (chapter 2) and the interpretations in social science research ethics (chapter 3). The conceptual framework is then assessed through its application to the first case study (CF2, chapter 4) and the second case study (CF3, chapter 5). These case studies are stand-alone chapters that aim to answer both the second and third research questions. This thesis assesses case studies in a sequence of increasing technical complexity to allow for a logical progression of the analysis. Finally, to answer the main research question, the development of the conceptual framework is then analysed in greater depth in the literature-based scrutiny (Chapter 6).

Case selection criteria

This thesis uses two case studies that differ in their data collection method and technical complexity. The criteria for selection are as follows:

Criterion 1

The selected research studies must process data collected from individual Internet users and thus be relevant for a research ethics evaluation by the Common Rule.

Criterion 2

The subjects must not be aware that data generated by them or their devices is being used for the technical Internet research analysed in the case study.

Criterion 3

One case study processes *passively* collected data and the other case study processes *actively* collected data.

Criterion 4

The case studies include a section addressing research ethics from which it is possible to discern their justifications and decision-making processes.

Chosen cases

The first case study is a paper titled “The Next Domino to Fall: Empirical Analysis of User Passwords across Online Services” by Chun Wang, Steve T.K. Jan, Hang Hu, and Gang Wang at the Department of Computer Science of the Virginia Polytechnic Institute and State University (popularly known as Virginia Tech). In this study, the researchers used millions of username and password combinations that were hacked

from popular online services and leaked to online forums. The data is comprised of real usernames and passwords of Internet users, none of whom were aware that their data would be used for a research study. This data collection is therefore *passive* because the researchers did not set up a new data collection mechanism to generate primary data. Instead, they reused existing data for a new purpose. The paper's section on 'Ethic Guidelines' speaks to how the researchers attempted to comply with the Common Rule.

The second case study is a paper titled "Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests" by Sam Burnett and Nick Feamster at the Georgia Institute of Technology and Princeton University respectively. These researchers measured potentially censored websites across the globe and sent the measurement results back to their own server. They then developed an invisible (controversial, though innovative) practice of initiating censorship measurements using people's devices without their knowledge when these users visited specific websites. This data collection is *active* because the researchers devised the data generation method and created information flows that would otherwise not have existed. The paper's extensive section titled 'Ethics and Security', as well as the public responses of IRBs, programme committees and other scholars to the research ethics concerns are instructive for this thesis.

Method of analysis

To determine how technical Internet research ought to be viewed from the perspective of research ethics, this thesis assesses the concepts of the Common Rule primarily through the lens of the ethics of care. Each part of the structured case

methodology has a different approach to uncovering, analysing, and further developing the concepts. This section differentiates between the first conceptual framework, the two case studies, and the final literature-based scrutinisation of the findings.

Initial conceptual framework

The first conceptual framework that forms the basis of this thesis is comprised of two separate chapters: Chapter 2 'Historical Emergence of Research Ethics' and chapter 3 'Conceptual Development of Internet Research Ethics'. These chapters, their function, and their focus are briefly introduced here. The initial conceptual framework (ie. CF1 in Figure 2, above) is summarised in a table at the end of chapter 3.

Chapter 2 describes the historical development of research ethics, its legal and philosophical foundations, and its codification in law of soft-regulatory instruments. The chapter then focuses on the specific concepts included in the US Common Rule, illustrating how the concept has been interpreted and critiqued in academic literature. This historical and legislative overview serves to set the baseline to assess criticism, commentary, and clarifications from academic literature with regards to research ethics law.

Chapter 3 develops the first iteration of the conceptual framework by building on the relevant concepts of chapter 2. The chapter takes a dynamic sociotechnical lens, whereby Internet technology has significantly impacted the understanding of information, research, and human interactions. The Internet has grown significantly over the last few decades, connecting more people via a variety of platforms and

services. Technical Internet research increased to help understand this growing technical phenomenon. Further, as daily human life became more integrated with and mediated by the Internet, social science research using the Internet as a venue also increased in volume and contributors.

The initial conceptual framework is developed as follows: First, the chapter describes how information laws (such as Research Ethics laws) are challenged by and reinterpreted in the new digital environments, to create a baseline understanding of why these concepts must be scrutinised for their relevance. Then, the chapter offers an overview of how scholars have responded to ethics review and applied the relevant concepts to studies conducted on or via the Internet. Social science and technical Internet research are separated into two sections. It describes the development of Internet research ethics in the two fields to date. Finally, the chapter constructs an initial conceptual framework of the current understanding of how the concepts have been interpreted in research ethics generally and for the Common Rule specifically. The initial framework is drawn up in tables, which are further developed throughout the case study chapters.

Case studies

The case study analyses are split into four parts, to reflect the various stages and methods of analysis. The first section of each case study situates the paper in the context of the intellectual environment to which it contributes. While it may read as a literature review, the aim is to illustrate how researchers in these fields consider the risks and benefits of their type of research, how they discuss and justify their work within research ethics review frameworks, and which technological options they have at their disposal for exploring alternative research designs.

The second part of each case study chapter answers the second research question by determining which concepts of the Common Rule are challenged in the pertinent paper. In each case study, the thesis recreates and explores the decision-making and reasoning methods likely used by the researchers and IRB. The decision charts provided by the US Department of Health and Human Services are helpful for ensuring that the order of considerations stipulated by the law are followed. To ensure the decision-making process is recreated correctly, I collected as much information about the procedure as possible. For example, the papers' sections on research ethics, correspondence with IRBs and programme committees, and any published information helps to clarify how the IRB came to its conclusion. This assessment allows me to determine which specific concepts of the Common Rule are too ambiguous to yield a clear decision, where the binary nature of the prescribed decision procedure offered too little room for interpretation, and which concepts were inappropriate for the particular type of study. The third part of the case study chapters answers the third research question, developing the conceptual framework further. Each concept identified in the previous section is deconstructed and reassessed through various disciplinary lenses of care. This part constitutes the iterative intellectual contribution of the thesis, which culminates in the final analytical chapter. The fourth part of the case study chapters develops the tables of the conceptual framework further. These are added at the end of each chapter to serve as an overview of the previous chapter and the developing conceptual framework.

Literature-based scrutiny

The sixth chapter of this thesis presents a literature-based scrutiny of the iterative conceptual framework that is developed throughout the case study chapters. The chapter answers the first research question, assessing the research ethics review procedure for technical Internet research, specifically through the lens of the ethics of care. This chapter proposes a new interpretation of the various concepts that were deemed to be unclear. Further, the chapter clearly articulates the meaning of these findings for the researchers, institutional review boards, policymakers, and the philosophical community engaged with the ethics of care.

Layout of the thesis

Chapter 2: Historical Emergence of Research Ethics

The second chapter provides the legal and historical basis for this socio-legal study. This chapter explores the origins, rationale, and history of research ethics. It provides an outline of the multiple approaches to research ethics in a variety of fields and describes the approach to human subject research in the US and Europe. It focuses specifically on the developments that led to the establishment of the Common Rule, the US human subjects-based research ethics framework codified into law in the Code of Federal Regulations for the Protection of Human Research Subjects (Title 45 CFR 46). This chapter further describes how legislators and scholars conceptualise the various concepts, values, and obligations that guide research ethics review.

Chapter 3: Conceptual Development of Internet Research Ethics

Social scientists applied the human subjects research model of research ethics to their research on Internet users before the technical Internet research community began to consider themselves implicated within this framework. This chapter establishes the conceptual framework for the case study analyses by compiling the lessons learned

in social science and the nascent discussions in the technical Internet research community. This chapter describes controversies and common points of ambiguity that confronted social scientists over more than two decades of Internet research and scholarly reflection. In doing so, this chapter reflects on the fundamental paradigm shifts caused by the Internet regarding assumptions about the concept of information (from a static resource, fixed to a carrier, to a dynamic informational environment in which data is easier to reuse) and the relationship between researcher and data subject. This chapter further illuminates the emerging debate about research ethics review with human subjects in the technical Internet research community. As such, it provides a basis for the interpretation of research ethics concepts, which will be tested, reviewed, and strengthened over the course of this thesis.

Chapter 4: Case Study 1 — Password Dumps in Research

The first case study assesses a paper that used more than 28 million individual password and username combinations that were hacked or leaked from Internet services. These databases, colloquially known as ‘password dumps’, are readily available on various websites and forums and have been used for many technical research studies. However, there is little consensus within the computer science community about the research ethics of using password dumps for research, nor is there clear guidance on taking research data from illicit sources. A close interrogation of the claims made by the researchers and the decision-making procedure of their institutional review board reveals that the concepts of harm and the impracticality of informed consent must be reviewed and reinterpreted. Furthermore, the case reveals the usefulness of applying the ethics of care, rather than the traditional cost-benefit or rights-based approaches, to research ethics.

Chapter 5: Case Study 2 — Internet Censorship Research

The second case study assesses a paper in which researchers proposed to use unsuspecting people's devices as vantage points to collect data on the censorship methodologies of controversial information from all over the world. In this case, the researchers would actively produce research data rather than relying on data that was hacked or leaked by other parties. This study was initially approved by the responsible institutional review board but met with criticism in the computer science community at the conference where the study was submitted. Examining the initial reasoning of the institutional review board and the subsequent decisions made by the conference programme committee reveals that several Common Rule concepts caused ambiguity and would benefit from further interpretation. These include the concepts of human subjects, at which point informed consent is required, the assessment of downstream risks of harm, autonomy as it relates to the concept of freedom as non-interference, and broader issues around transparency, accountability, and compliance with the law when a project is not bound to a specific jurisdiction but operates globally.

Chapter 6: Literature-based Scrutiny of Research Ethics Frameworks for Technical Internet Research

The ongoing analysis in the two case studies culminates in the analysis chapter. Using the case study analyses, this chapter identifies the ambiguous places in the Common Rule and the inappropriate provisions in the law for technical Internet research. This chapter then collects the reinterpretations suggested in the cases and assesses them holistically, using the literature that informed chapter 2 and 3 as well as that which became relevant in the case study analyses. Finally, the chapter

formulates a new legal provision of the concepts that were found to be ambiguous. The new legal provisions are underpinned by a qualitative assessment of technical Internet research and a more appropriate ethical framework.

Chapter 7: Conclusion

The final chapter concludes the thesis. It reflects on the findings, suggests practical contributions for affected groups (e.g. technical Internet researchers, research ethics committees, legislators or research ethics policy makers, and the ethics of care community in academic philosophy). It further suggests new avenues for further research, which were beyond the scope of this thesis.

(2) Historical Emergence of Research

Ethics

Introduction

The human subject model of research ethics review is an attempt to formalise the responsibilities of researchers with regards to their data subjects. The model has been implemented in a variety of frameworks, ranging from formal laws, to disciplinary covenants, and departmental guidance. This thesis tests the appropriateness for the review of technical Internet research of the US legal implementation of the human subjects model in the so-called Common Rule, enacted in the U.S. Department of Health and Human Services' Title 45 CFR 46 (Public Welfare) Subparts A, B, C and D.⁸ The aim of this chapter is to lay the legal, philosophical, and practical foundations of this thesis. A deep socio-legal inquiry into research ethics law and policy necessarily requires a baseline description of several concepts, their origins, and their interactions. To this end, the chapter is divided into four sections to set the scene for this thesis.

Part 1 of this chapter positions the role of researchers in an interdependent society in terms of their relationships, dependencies, power, and resulting responsibilities. It describes the justification and limitations of academic freedom as a foundational principle of the profession and the role of research ethics to uphold standards. Specifically, the relationships of researchers are described, which inform the responsibilities a researcher must observe.

⁸ Available at: <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html>.

Part 2 describes the historical emergence of research ethics review models as a response to atrocities committed in the name of research. Specifically, the section traces the emergence of the human subject model of research ethics review. This model was first formally codified as a response to the atrocities of the Second World War. Today, several international conventions stipulate the baseline considerations for the human subject research ethics model. The adoption of these conventions differ around the world, per academic discipline, per university, and per department. The US interpretation of the human subjects model in a formal, national law is not necessarily common in the world, and therefore a relevant subject of study. This section also explores Europe's approaches to the human subject model of research ethics and describes diverging disciplinary approaches.

Part 3 examines the various concepts that constitute the research ethics review procedure in the US Common Rule. The main focus is on the concepts of human subjects, the risk of harm, autonomy, and privacy. This part of the chapter situates the concepts within the Common Rule procedure, discusses their ethical and philosophical underpinnings, and engages with their interpretations in scholarly discourse.

Finally, part 4 describes the interplay between law and ethics, and specifies the research themes of this thesis. The ethical responsibilities of researchers are, unsurprisingly, interpreted through various ethical lenses and then codified into some form of law (be it formal or informal). This part therefore necessarily but briefly introduces the various ethical theories that are applied throughout this thesis.

Further, it describes the type interpretation of research ethics law as communicative of ethical values and procedures.

Part 1: Researchers' Roles, Relationships, and Responsibilities

Academic researchers have a multi-faceted position in society. They are expected to contribute to knowledge by conducting research and teaching. Knowledge creation and dissemination is assumed to be useful to society in a variety of ways, for example by stimulating innovation, informing policy-making, and empowering individuals with new knowledge and insights. The work of the academic researcher is—in part—enabled by the professional autonomy provided by the principle of *academic freedom*. However, some limitations emerge from the relationships within which academic researchers necessarily operate. This section positions the researcher and locates the source of their responsibilities.

Primarily, academics are considered to be experts on specific topics (Aby & Kuhn, 2000). They are mostly free to choose their topic of expertise, their research methodologies, and the way in which their knowledge is disseminated through teaching and publication. The freedom and autonomy to conduct research and teaching are dependent on the principle of academic freedom (Russell, 2002). The scope of this principle has become a defining feature of academic researchers' professional identities.

In theory, academic freedom facilitates the higher-order value of *knowledge production*. In this light, Conrad Russell referred to academic freedom as enabling the

‘pursuit of knowledge for its own sake’ (Russell, 2002). However, beyond the production of purely theoretical knowledge, the creation of new knowledge can also serve as both an informed, deliberative democratic process and an innovative economy. For example, a democracy is considered to thrive when it allows new ideas to be discussed and its policies to be based on evidence. Furthermore, economic progress can be stimulated when the fruits of academic research, such as new ideas and inventions, are produced and commercialised.

The principle of academic freedom is today adopted into many nations’ constitutions and in various international declarations as a fundamental human right.⁹ The current understanding is that universities have freedom within the framework of the law.¹⁰ In the UK, the principle of academic freedom is legislated as follows:

[...] to question and test received wisdom and to put forward new ideas and controversial or unpopular opinions without placing themselves in jeopardy of losing their jobs or the privileges they may have.¹¹

The scope and interpretation of academic freedom changed over the centuries, and was gradually limited in response to immoral activities or atrocities within the domain of science. Several centuries ago, universities were initially founded with the aim of allowing the free and unimpeded exchange of ideas through teaching, research, and deliberation. The ideal of freedom in this regard can be traced back to

⁹ See for example the UNESCO’s “Recommendation concerning the Status of Higher-Education Teaching Personnel”: http://www.unesco.org/education/docs/recom_e.html

¹⁰ The University of Oxford shut down from the years 1209 to 1214 as a result of a quarrel between the university and the town’s Mayor. The university claimed to have jurisdiction over an internal murder case, while the legislator argued that this crime does not fall within the scope of the institution’s autonomy. The resulting settlement laid the foundation for the scope of academic freedom. See <https://twitter.com/bodleianlibs/status/1141722670234554368>

¹¹ See Part IV, article 202(2)(a): <http://www.educationengland.org.uk/documents/acts/1988-education-reform-act.pdf>.

the autonomous position of churches in medieval times, when universities and the Church were closely aligned or even intertwined (Boele-Woelki, Fisher, Pape, & Francisco, 2019). The first precedent against absolute academic freedom was set in the 13th century, when the Mayor of Oxford ruled that the university did not have its own jurisdiction over a particular murder case. This decision led to the constitution and charter of liberties of the university.¹² Scholars struggled to maintain the autonomy of university institutions over many centuries (Aby & Kuhn, 2000). One of the most recent limitations of academic freedom is the requirement for standardised research ethics review. Following a range of atrocities in the name of science throughout the 20th century, legal frameworks developed to stipulate how an ethical review of research must be conducted.

Limitations on academic freedom originate from—and are legitimised by—the balancing and corrections of power academics can exert within their relationships (as described in the next paragraph). Lis Lange, paraphrasing Pierre Bourdieu, expressed this convincingly when she stated that:

[...] it is in our liberties as academics where our responsibilities are located. These liberties are simultaneously exercised in relation to our academic practice and in relation to the society at large. (Bourdieu, 1988; Lange, 2016, p. 184)

Researchers are embedded in a variety of social environments and relationships that arise naturally from their activities. These relationships are typically with data subjects in experimentation or research studies, their colleagues in their institution or

¹² See <https://www.admin.ox.ac.uk/statutes/375-092.shtml>. See the original document here: <https://twitter.com/bodleianlibs/status/1009392890383040512>.

the wider research profession, and their dependency on society's patronage embodied by political representatives and government institutions. Some of these relational contexts may be characterised by power imbalances, dependencies, and hierarchies.

Researchers typically work directly with research participants to generate and collect data. They may also use secondary data sources, in which they do not have a direct relationship with their data subjects, but a somewhat indirect relationship.

Researchers invariably have more knowledge about the experimental procedures and their purposes than their data subjects. Therefore, the researchers yield some form of power over their data subjects by collecting data about them, or with the ability to change a data subject's environment in the pursuit of knowledge (LaRossa, Bennett, & Gelles, 1981; Meara & Schmidt, 1991).

Katarina Boele-Woelki explains that academic freedom is tolerated by society and its public opinion, based on society's trust that academic institutions are generally useful and serve legitimate ends (Boele-Woelki et al., 2019). Indeed, the work of researchers is typically funded by governments, and thus their livelihood is directly dependent on the will of policymakers to maintain research institutions (Russell, 2002). The resources that governments can allocate to research activity are typically raised through tax collection. Therefore, researchers have an indirect relationship with the societies in which they are embedded, yet are responsive to public opinion to justify their existence. Society's willingness to support research—which is based on public opinion and public trust—is the pillar that supports academic research and knowledge production. Moreover, policymakers make use of academic output in

their deliberations to set the course of societies and allocate resources, from which a further social responsibility for scholars arises.

Directly related to the importance of social and political trust to maintain a steady flow of funding, researchers also have a relationship with their fellow researchers and the research profession as a whole, beyond mere collegial ties. If a researcher violates public trust with his or her work, the reputation of the academic profession may be tarnished. This may result in a decrease in funding or in further, stronger limitations on the principles of academic freedom.

The relationships described above show varying degrees of power and dependency. Researchers are initially powerful as specialists in particular areas of knowledge who may conduct experiments and collect data in the service of the higher goal of knowledge production. Data subjects, politicians, and institutions have a certain degree of trust in the way researchers conduct their work and publish their findings. However, politicians and institutions maintain some power over researchers in their patronage relationship. Further, data subjects, as human research participants who are members of society, may sway public opinion against the research profession, if researchers yield their power in immoral or irresponsible ways. Finally, researchers have an interest to police their colleagues in their disciplines, their institutions, and the wider research profession, to maintain the required level of trust from society, politicians, and institutions.

Academic freedom must be understood within this web of relationships¹³ and resulting responsibilities. It is not an absolute principle, but instead rests on a tightly

¹³ The term “web of relationships” is derived from Carol Gilligan’s work titled *In a Different Voice: Women's conceptions of self and of morality* (Gilligan, 1977) and the later interpretations of this foundational work.

knit foundation of power, trust, and responsibility. To justify and maintain their autonomy, academics have an interest in maintaining their reputations by upholding scholarly integrity, taking action against fraud and abuse, and avoiding committing harm and atrocities. Over the centuries, this has required the imposition of rules and regulations by legislators or by academics themselves. The reliance on maintaining public trust should, in theory, instil in researchers a deep sense of social responsibility and public service.

The human subject model of research ethics focuses only on the responsibilities towards data subjects. It may indeed be argued that this is the researcher's primary relationship. However, a healthy relationship between researchers and their data subjects likely translates into good relationships with society and colleagues, too. The next section explores the historical emergence of this model of research ethics review.

Part 2: Models of Research Ethics

The questions regarding the right and wrong actions of researchers are as old as research endeavours themselves. In the 1st century AD, for example, a Roman physician named Aulus Cornelius Celsus wrote in his medical treatise, *De Medicina*, that 'it is not cruel to inflict on a few criminals sufferings which may benefit multitudes of innocent people through all centuries' (Celsus, 47 AD). Celsus thereby justified cruel and possibly torturous human experimentation on criminals in the name of benefiting humanity. It can be argued that he was reflecting the values of his time.

It is now commonly acknowledged that research produces values or goods such as knowledge and understanding, but it may also raise further questions, dilemmas, and unintended consequences (Owen et al., 2013). Since Celsus, research ethics has evolved to take into account not only scientific progress but also a concern for the rights and welfare of the research subjects. This section discusses some atrocities that have been perpetrated against human subjects in the name of research, and the responses from society, legislators, and the scientific community that guided the emergence of the human subject research ethics model.

Atrocities

Ethical lapses in research have a long history. This section describes a few of the best-documented examples that have contributed to an understanding of how research ought to be regulated. It is striking that the researchers in most of these examples could have justified their approach using Celsus' reasoning that it is acceptable to experiment on certain groups and thereby violate their rights and welfare for the benefit of humanity.

Atrocities in research have been well-documented in the US. In fact, the US legal framework regarding research ethics is a direct regulatory response to the malevolent behaviour of researchers throughout the 20th century. A famous case is the Tuskegee syphilis experiment, which ran from the early 1930s until the 1970s. A consortium of US researchers purposely withheld treatment for syphilis on specific minority groups to study the disease's spread among that population group (Cobb, 1973). The intent was to understand ways to halt the spread of syphilis, though this knowledge was created at the expense of a large group who could have been treated instead.

Other examples in the US include the CIA's decades-long brainwashing and mind control experiments entitled Project MKUltra, which were halted in 1973 (Senate, 1977); the Stanford prison experiment on the perceived effects of authority and power relations in 1971 (Zimbardo, Haney, Banks, & Jaffe, 1972); and the Milgram experiments on obedience to authoritative figures (Milgram, 1963). Earlier, in the 1960s, healthy children at the Willowbrook State School in New York City were intentionally infected with hepatitis by feeding them solutions that included the faeces of other children who had been infected with the virus to monitor the effects of antibodies in combating the virus (Rice, 2008).

However, some of the most atrocious examples of research subjects being abused took place in Europe. Before and during the Second World War, doctors of the Nazi regime conducted a range of experiments on prisoners in concentration camps. These experiments included (but were not limited to) freezing prisoners to research hypothermia, immunisation for the treatment of malaria, experimentation with treatments for wounds caused by mustard gas and chemical agents or bombs, testing the effects of drugs, experimentation on effective human sterilisation and fertility treatment, and experimentation on adolescent twins for easily acquired access control groups (Shuster, 1997).

Reactionary Regulation: The Human Subject Model

The resulting public outrage to abuses such as these initiated public discussions about the skewed balance of individual harms and research benefits. These discussions led to several regulatory responses whereby legislators enacted sets of laws that set out the process by which research ethics would be safeguarded in future (Fairchild & Bayer, 1999).

Specifically, the human subject model of research ethics has been codified into a variety of laws that operate on institutional, national, and international levels. The first regulatory milestone for the human subject model of research ethics was the Nuremberg Code (Code, 1949). During the Nuremberg trials, the doctors of the Nazi party were held accountable and punished for the inhumane and murderous experiments conducted or overseen by them, which included their experimentation on their prisoners. The judges at the Nuremberg trials authored and published the Nuremberg Code, which provided a set of principles to operationalise ethical considerations in research. The judges did not concern themselves with distinctions in research fields (e.g., biomedical science, behavioural science, and social science), although the Nuremberg Code was clearly drafted in response to abuses of power in medical research by the Nazi doctors (Feldman, 2012). Provisions in the Nuremberg Code included mandatory informed consent, a favourable risk and benefit analysis, and the right for research subjects to withdraw without repercussions (Code, 1949; Rice, 2008). The Code's focus on human research participants led to the creation of the so-called human subject research model. While the Nuremberg Code was initially ignored by practising researchers or only considered to be a useful guideline rather than a binding document (Edelson, 2004), it became the foundation upon which other laws, treaties, and research ethics guidelines were based.

In the 1960s, the Nuremberg Code was used as the guiding document for the World Medical Association's Declaration of Helsinki, which was developed to be a 'statement of ethical principles for medical research involving human subjects, including research on identifiable human material and data' (Association, 2013). The Declaration has been amended nine times to date, most recently in 2013. It strengthens the principles of the Nuremberg Code and includes two new principles:

that the interests of the research subject should always be placed above the interests of society and that every subject should receive the best treatment available (Rice, 2008). However, the Declaration weakens the Nuremberg Code's absolute requirement for informed consent, appreciating that obtaining consent from subjects is not always possible. While the Declaration of Helsinki is not enforceable under international law, it has been labelled as the *de facto* constitution for medical sciences and bioethics (Benatar & Singer, 2000; Rickham, 1964). Moreover, scholars have noted that the Declaration draws its authority from the extent to which it has been codified in, or has influenced, national laws around the world (Human & Fluss, 2003). These implementations have led to diverging processes for research ethics. The US and Europe, for example, have based their research ethics requirements on the Declaration of Helsinki (Association, 2013), although the resulting review procedures imposed on researchers are different.

Codification into Law

The sources of research ethics requirements are found in a range of international declarations, national laws, and guidelines or requirements in specific disciplines provided by overarching organisations, universities, departments, and funding organisations. Furthermore, academic journals or conferences may impose action-specific guidelines or may leave the ethical judgement to a panel of academic peers. While some research ethics requirements are rooted in law, others may be equally relevant through enforcement by publication venues or funding organisations.

This section illustrates some of the principles and procedures required in the US and EU research ethics frameworks. However, scholars have noted that the US and EU have distinctly different approaches to research ethics. US research ethics law and committees focus the review on the favourable outcome of a risk and benefit

analysis, only requiring research to minimise any harm to research subjects. Buchanan and Ess identified this as a utilitarian or consequentialist approach, whereby the net benefit to society is decisive (Buchanan, & Ess, 2008). In contrast, European approaches to research ethics review focus more on the rights of individual research subjects, which is a deontological approach (Ess, 2006).

The United States

As part of its response to atrocities in research, the Federal Government of the US enacted the US National Research Act in 1974, legislation that introduced a requirement for research oversight in the form of IRBs at US-based research institutions. To guide the operation of IRBs, the Act established the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. This commission published its guidelines, the Belmont Report, in 1978 (Commission, 1978). The Belmont Report identified three fundamental ethical principles through which research projects should be assessed: respect for persons, beneficence, and justice.

The United States Department of Health and Human Services adopted the Belmont Report's findings in a policy called the Common Rule, which was later enacted into legislation in 1991 under the Code of Federal Regulations for the Protection of Human Research Subjects (Title 45 CFR 46) in its subparts A, B, C, and D (Services, 2009). Today, the Common Rule is the baseline standard of research ethics by which government-funded research in the US is evaluated.

The task of IRBs is to protect human data subjects and serve as their advocates in the research ethics review process, based on a review of the methods and technologies proposed (Aycock, Buchanan, Dexter, & Dittrich, 2011; Pittenger, 2003). Beyond

protection of the subjects, it has been argued that reviews of studies should improve the researchers' levels of ethical literacy and encourage them to reflect on their project from the perspective of all involved or affected stakeholders (Wiles, & Boddy, 2013). The operation and considerations of an IRB are carefully outlined in the Common Rule, which provides several criteria and exemptions for research ethics review that an IRB is obliged to consider in the review of a proposed study. The US Office for Human Research Protections has published several decision trees to support an IRB's decision-making.¹⁴ The relevant requirements and their definitions are discussed in the next sections.

European Union

The European Union (EU) and its member states have regulated human subject research ethics less formally than the US. Existing research ethics regulation exists on several levels: on the supra-national level (in the EU) and on the national level (in individual member states). This section discusses the two levels in turn.

Two formal legislative documents of the EU address research ethics directly, albeit sparsely. Firstly, the EU adopted a Clinical Trials Directive (2001/20/EC), which focuses on medical research involving human subjects. Article 6 of the directive requires EU member states to establish ethics committees for clinical trials. The task of these committees is to protect the rights, safety, and well-being of human subjects involved in trials, and to provide its opinion on the proposed trials. EU member states have implemented ethics committees for clinical trials in a variety of ways, with varying degrees of power over research (Hedgecoe, Carvalho, F., Lobmayer, P., & Raka, F, 2006). However, no specific or standardised guidance has been provided

¹⁴ See <https://www.hhs.gov/ohrp/regulations-and-policy/decision-charts/index.html>.

on how the ethics committees should evaluate research proposals, or which specific ethical principles should be considered in the review process. To illustrate why this is problematic, the lack of standardised guidance has led scholars from genomics and medical sciences to complain that the patchwork of terminology, legal standards, and research ethics guidance prohibits the sharing of biobanks, which reduces their utility and has created barriers to beneficial, cross-border research and international research consortiums (Colledge, Elger, & Howard, 2013; Elger & Caplan, 2006; Knoppers & Saginur, 2005; Zika, Schulte In den Bäumen, Kaye, Brand, & Ibarreta, 2008).

Secondly, the Horizon 2020 research programme—Europe’s largest funding programme for research and innovation to date—was established with explicit reference to ethical principles in article 19 of its founding Regulation (No. 1291 / 2013):

All the research and innovation activities carried out under Horizon 2020 shall comply with ethical principles and relevant national, Union and international legislation [...]

Furthermore, article 13 of the Regulation makes adherence to ethical principles more enforceable:

A proposal which contravenes ethical principles or any applicable legislation [...] may be excluded from the evaluation, selection and award procedures at any time.

Finally, article 14 further states that the European Commission will

[...] systematically carry out ethics reviews for proposals raising ethical issues. Such reviews shall verify the respect of ethical principles and legislation [...]

Documents relation to the Horizon 2020 programme refer to three separate guidelines for researchers. The first is a 41-page checklist document entitled 'Horizon 2020 Programme – Guidance: How to complete your ethics self-assessment'¹⁵. The topics covered include human embryos, human beings and their personal data, animals, the environment, and the potential future misuses of research findings. The document states that its aim is merely to the researcher's "[...] proposal 'ethics-ready' for Horizon 2020 funding", but does not provide a mechanism for a standardised ethics review.

In addition, the EU has published several guidelines that give non-binding advice to researchers about research ethics that also provide useful definitions. The most detailed publication in this regard is the European Textbook on Ethics in Research, published in 2010 to support the Seventh Framework Programme (which is the precursor to the Horizon 2020 programme) (Hughes, Hunter, Sheehan, Wilkinson, & Wrigley, 2010). While the book was authored by persons trained in bioethics and the examples used in the book are mostly related to medical or behavioural sciences, the concepts explored in the book are relevant beyond these disciplines. To a certain extent, the book follows the structure of the US Belmont Report and its principle-based approach; for example, the book discusses the importance of consent and the duty of care towards vulnerable people as an equivalent to *respect for persons*.

¹⁵ See <http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf>.

Similarly, the book discusses the balancing of risks and harms as an equivalent to *beneficence*, and it devotes a chapter to *justice* in research. As a result, many of the definitions used in the Textbook are similar to those used in the US. For example, research is defined as having the purpose '[...] generate (new) information, knowledge, understanding, or some other relevant cognitive good, and does so by means of a systematic investigation.' (Hughes et al., 2010). Another document, The European Code of Conduct for Research Integrity, provides some high-level principles that are applicable to the profession of researchers (e.g., an overview of good research practices, such as data management, collaborative working, and dealing with misconduct). However, it only instructs researchers to comply with codes and regulations that are relevant for their disciplines, without naming them.

Finally, Horizon 2020 documents frequently refer to the concept of 'Responsible Research and Innovation' (RRI). The aim of RRI is to give a voice to stakeholders who may be involved or affected by a research project at an early stage (van den Hoven, 2013). The RRI concept is described as an evolving framework that establishes a 'higher-level responsibility or meta-responsibility that aims to shape, maintain, develop, coordinate and align existing and novel research and innovation-related processes, actors and responsibilities with a view to ensuring desirable and acceptable research outcomes' (Stahl, 2013). RRI is influenced by several disciplines, among which technology assessment, applied ethics, and science and technology studies research are the most notable (Grunwald, 2011).

It is also notable that none of the guidelines referenced by the EU define the concept of *human subjects* directly, although some documents use the term. Other documents use terms such as 'human participants' and refer to the Declaration of Helsinki.

Finally, the guideline documents published within the EU typically require researchers to comply with existing laws (for example, see (KNAW, 2016)) but mostly refer to the Data Protection Directive (95/46/EC) and its successor, the General Data Protection Regulation (Council of the European Union, 2016).

Most of the EU member states' national laws on research ethics specifically concern the biomedical sciences and clinical trials. For example, the UK has implemented a law called the Medicines for Human Use (Clinical Trials) Regulations 2004. The UK's Economic and Social Research Council has provided a guideline document entitled 'The Research Ethics Guidebook: A Resource for Social Scientists', in which it refers to the following laws as guiding researchers' behaviour: the Data Protection Act 1998, the Human Rights Act 1998, the Mental Capacity Act 2005, the National Health Service Act 2006, the Protection of Vulnerable Groups (Scotland) Act 2007, the Safeguarding Vulnerable Groups Act 2006 and the Adults with Incapacity (Scotland) Act 2008 (Boddy et al., 2010).

Similarly, a recent study into research ethics by the Dutch Royal National Academy of Sciences found that no formal laws exist to prescribe research ethics review beyond the medical sciences (KNAW, 2016). The Dutch law on higher education and scientific research (WHW)¹⁶ states in its article 1.6 that academic freedom is to be observed by research institutes. Article 1.7 of this law grants governing bodies a mandate to develop their own guidelines with regard to research ethics review and governance.

¹⁶ See <<https://wetten.overheid.nl/BWBR0005682/2019-01-01>>.

A survey of EU member state regulations concerning research ethics found that most states regulate biomedical research to varying degrees (Rangi & Warso, 2015).

Beyond that, however, the examples of the UK and the Netherlands appear to be typical of the state of regulating human subject research in Europe. Rangi and Warso's survey report notes that despite the lack of human subject research regulations in EU member states, 'there are various methods addressing the topic, most often codes of conduct, as well as the general non-legislative, such as institutional, procedures governing ethics assessment of research involving humans.'

Overview of Formal Regulatory Structure

The below table lists the various regulatory and policy documents that have been developed specifically for human subject research or that have been described as relevant by governments, academic institutions, and scholars. The table names the regulatory document, its authoring institution, its jurisdiction, the academic discipline of research for which it is stated to be relevant, and whether its obligations are mandatory for researchers. For the sake of completeness, the table contains more regulatory and policy documents than are described in the previous sub-sections. The documents that have not been previously described are not historically significant to the development of human subject research ethics or are overarching international laws that are beyond the scope of this thesis.

Name of policy	Organisation	Formal jurisdiction	Discipline	Mandatory?
Nuremberg Code ¹⁷	Nuremberg trial judges	None	Research broadly developed in response to	No

¹⁷ <<https://history.nih.gov/research/downloads/nuremberg.pdf>>.

			atrocities committed in medical research	
Belmont Report ¹⁸	National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research	USA	Biomedical and behavioural sciences	No
Title 45 CFR 46 ¹⁹	United States Department of Health and Human Services	USA	Human subject research	Mandatory
Declaration of Helsinki ²⁰	World Medical Association	None	Biomedical research	No
Universal Declaration on Bioethics and Human Rights ²¹	UNESCO	Universal	Medicine, life sciences and associated technologies as applied to human beings	No
International Ethical Guidelines for Biomedical Research Involving Human Subjects ²²	Council for International Organizations of Medical Sciences	None	Biomedical Research Involving Human Subjects	No
Oviedo Convention on Human Rights and Biomedicine ²³	Council of Europe	Signatory countries	Applications of Biology and Medicine	Yes
Clinical trials - Directive 2001/20/EC ²⁴	EU	EU Member States	Clinical trials on medicinal products for human use	Yes
Horizon 2020 ²⁵ / Responsible Research and Innovation ²⁶	EU	EU Member States	All research proposals	Yes

¹⁸ <<https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>>.

¹⁹ <<https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html>>.

²⁰ <<https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>>.

²¹ <<http://www.unesco.org/new/en/social-and-human-sciences/themes/bioethics/bioethics-and-human-rights/>>.

²² <<https://cioms.ch/shop/product/international-ethical-guidelines-for-biomedical-research-involving-human-subjects-2/>>.

²³ <<https://www.coe.int/en/web/bioethics/oviedo-convention>>.

²⁴ <https://ec.europa.eu/health/human-use/clinical-trials/directive_en>.

²⁵ <http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/ethics_en.htm>.

²⁶ <<https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation>>.

EU Member State Biomedical Research Law	National governments	EU Member States, individually	Biomedical research	Yes, to the extent compliance with individual laws has been made mandatory
EU Charter of Fundamental Rights ²⁷	EU	EU Member States	All disciplines	Yes
European Convention on Human Rights ²⁸	Council of Europe	Signatory countries	All disciplines	Yes
General Data Protection Directive ²⁹	EU	EU member states and all EU citizens' personal data	All disciplines	Yes

Figure 1. Overview of Formal Regulatory Structure

Other Models of Research Ethics

Research ethics does not focus solely on human subjects, and the human subject model is one of several research ethics frameworks. Other models of research ethics frameworks exist in the humanities, the natural sciences, and engineering, for example. This section briefly outlines the shift to the social sciences and presents some other approaches to research ethics from other disciplines.

Social Sciences

Research in the social sciences was subjected to formal ethics review in the 1960s and 1970s following a range of scandals and atrocities. The human subject model applied in biomedical science was transposed to the social sciences because both disciplines concern the study of humans and collect data from them. However, this imposition

²⁷ <https://ec.europa.eu/info/aid-development-cooperation-fundamental-rights/your-rights-eu/eu-charter-fundamental-rights_en>.

²⁸ <https://www.echr.coe.int/Documents/Convention_ENG.pdf>.

²⁹ <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679>>.

of an ethical framework from one discipline to another has been met with criticism and resistance (Israel & Hay, 2006). For instance, Zachery M. Schrag stated that ethics committees impose “silly restrictions” on researchers and that “ethics review is a solution in search of a problem” (Schrag, 2011, p. 123). In his annual lecture to the Academy of Social Sciences in Australia, Philip Pettit reflected on the evolution of research ethics frameworks for social sciences and concluded that their imposition will “[...] lead to a serious reduction in the current scope of research and to a substantial compromise of the ethic that currently governs research practice” (Pettit, 1991, p. 100). Agata Gurzawska et al. found that “[...] the application on biomedically-based procedures has been met with considerable resistance, since the nature of ethical issues in social sciences is different from those in medicine” (Gurzawska & Benčin, 2015, p. 3).

Humanities

The humanities comprise academic disciplines in a wide range of fields, including language, literature, history, philosophy, jurisprudence, and the arts. The blanket imposition of the human subject research model on the humanities has been rejected by several research councils and other authoritative bodies. Instead, research ethics guidelines in the humanities focus more on values of scientific professionalism and specifically the value of preservation (e.g., of cultural artefacts and monuments). To the extent that the humanities collect data from human subjects, scholars recommend following the values and principles of the human subject model, although they note that the risk of harm to humans will typically be non-existent in the humanities in comparison to biomedical science or even the social sciences (Benčin, Šumič-Riha, & Riha, 2015).

Natural Sciences

Ethics guidance in the natural sciences, such as physics and chemistry, focuses on scientific integrity and the professional conduct of scientists. In addition, guideline documents describe the responsibility that scientists have towards society in advocating for the advancement of human welfare (Davis, 2002) and warning society of the harmful effects of their findings.³⁰ To the extent that humans are directly implicated in their research, researchers in the natural sciences are expected to follow general health and safety guidelines and to respect privacy, where applicable (Strle, Benčin, Šumič-Riha, & Riha, 2015).

Engineering

The field of engineering studies the design, modelling, and development of machines, materials, and processes for practical ends. Engineering ethics is mostly concerned with the professional roles and responsibilities of engineers and their scientific integrity. Due to the focus on the practical ends of research, it has been noted that engineering ethics is consequentialist in nature, concerned with the resulting effects of inventions (for example, see (Lloyd & Busby, 2003)). Engineering ethics differs from technology ethics, which studies the normative ethical issues that emerge from the introduction of technology into society.

Part 3: Concepts of the Common Rule

This section examines the concepts that constitute the US Common Rule. The law imposes a binary decision procedure for research review, whereby a yes/no answer

³⁰ See for example: Royal Society of Chemistry, *Regulation of the profession and code of conduct*, http://www.rsc.org/images/code-of-conduct_tcm18-5101.pdf

is sought to the question of whether the research study fits into prescribed categories and concepts. To illustrate, a review procedure typically needs to assess the following: research proposals must be reviewed by an IRB if they are, in fact, **research**, and collect **private information** about **living and identifiable human subjects** through an **intervention** or **interaction**. In most cases, the researcher must ask for **informed consent** from the data subjects. The review procedure may be expedited and the **informed consent** requirement waived for a number of reasons, the most relevant of which are when the study does not present **greater than minimal risk** to the human subjects, or when the **information collected already exists** and is **publicly available**. The words in bold are the important threshold criteria for assessing research proposals in the US, as prescribed by the Code of Federal Regulations for the Protection of Human Research Subjects (Title 45 CFR 46) in sections §46.101, §46.102, §46.110, and §46.111.

This part assesses these concepts in greater depth, synthesising the research ethics literature, legal analyses, and their philosophical underpinnings. Priority is given to the concepts that are the most relevant to this thesis: the definition of research, how human subjects are conceptualised, risk assessments and their balancing with foreseen benefits, the idea of enabling autonomy through informed consent and diminished autonomy in vulnerable populations, and the broad concept of privacy.

Research

US law defines research specifically in §46.102 as follows:

[...] a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge [...]

Snider and Stroup note that this definition appears to have been developed in response to atrocities in medical research, and may thus not necessarily be interpreted easily in other contexts or disciplines (Snider Jr, & Stroup, 1997, p. 30). The Belmont Report anticipated definitional challenges and proposed a limit to the definition, although still aimed at the biomedical field, when it stated that:

Even when a procedure applied in practice may benefit some other person, it remains an intervention designed to enhance the well-being of a particular individual or groups of individuals; thus, it is practice and need not be reviewed as research. (Commission, 1978, p. 10)

This limitation allows differentiation between research and other practices for which a researcher may collect information. Snider and Stroup encouraged IRBs to question the intent of a given data collection activity, which can be uncovered by assessing its design. If the primary intended purpose of the activity is to develop or contribute to generalisable knowledge, the activity should be considered 'research' for the purpose of IRB review. Pritchard reaffirmed this interpretation when he stated that the intention to contribute to generalisable knowledge can be found in the design of a data collection proposal. However, Pritchard disagreed with Snider and Stroup, who argued that the Common Rule should not apply when research is not the primary aim of a data collection, even though research is conducted on the collected data at a later stage. In Pritchard's experience, it is common for research to be influenced by practical work, and this secondary intent should not, therefore, be exempted from research ethics review (Pritchard, 2001).

Furthermore, US research ethics laws do not require a study to undergo research ethics review when that study is not federally funded. However, most—if not all—research institutions prescribe research ethics review for all studies, without discriminating between the sources of funding.

Human Subjects

Once it has been established that the proposed activity is indeed *research*, the Common Rule questions whether the designed research proposal involves *human subjects*. This concept is the crucial trigger to initiate a full IRB review. Human subjects are defined in §46.102(f) as follows:

[...] a **living individual** about whom an investigator (whether professional or student) conducting research obtains

- (1) Data through **intervention** or **interaction** with the individual, or
- (2) **Identifiable private information**.

To date, academic literature appears to have considered the concept of ‘living individuals about whom information is collected’ to be self-explanatory. A data point in a research project is either information about a living person, and thus a human subject, or it is not. The work by Carl Öhman may challenge this conception, as he continues to explore the ethical dimensions of conducting research and commercial exploitation on the data relating to deceased persons (Öhman, C. & Floridi, 2017; Öhman, C. & Floridi, 2018).

Similarly, little academic literature has considered the concepts of *intervention* and *interaction* with individual data subjects. The definitions provided in §46.102(f) of the

Common Rule are somewhat colloquial and contain examples of clauses that are only relevant to the medical sciences:

Intervention includes both physical procedures by which data are gathered (for example, venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes. **Interaction** includes communication or interpersonal contact between investigator and subject.

These definitions appear to be directed towards the clinical collection of medical data, both physically and through communication between the researcher and the patient or data subject. In the context of internet research, wording such as 'manipulation of the subject or the subject's environment' and 'communication' may be interpreted significantly differently than intended by the authors, which is explored in the case studies of this thesis.

The concept of 'private information' that is 'individually identifiable' can be found in the rich academic discourse on information privacy, although less so in the literature on research ethics. The Common Rule defines this concept in §46.102(f) as follows:

Private information includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record). Private information must be **individually identifiable** (i.e., the identity of the subject is or may readily be ascertained by the investigator or associated with the information) in order for obtaining the information to constitute research involving human subjects.

These terms and their definitions are similar to definitions in various US privacy laws. For example, the ‘reasonable expectation of privacy’ is a dominant concept in legal threshold in US law and academic literature to decide whether a piece of information should be considered as private (Wilkins, 1987; Winn, 2009). The ‘individually identifiable’ criterion mirrors standard definitions in US privacy law and regulation, such as the Health Insurance Portability and Accountability Act (Prevention, 2003) and the authoritative National Institute of Standards and Technology’s (NIST) *Guide to Protecting the Confidentiality of Personally Identifiable Information*, in which it defines personally identifiable information as

Information which can be used to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother’s maiden name, etc. (McCallister, Grance, T., & Scarfone, K. A., 2010, p. 45).

Risk of Harm vs. Benefits

The Common Rule requires in §46.111(a)(1) that the risks to subjects are minimised. Furthermore, the law prescribes a weighing of risks and benefits in §46.111(a)(2), as follows:

Risks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result.

Part 4 of this chapter describes a variety of ethics reasoning methodologies. However, it is relevant to note that a provision like this one corresponds with the utilitarian tradition of US (and arguably, UK) research ethics approaches, where moral calculations are acceptable to justify some harms. In the more deontological European research ethics traditions, an understanding of the violation of individual rights would be equally important to obtain before assessing the effects of the research on the rights and interests of the subject. In the US tradition, a violation of rights is part of the weighing exercise. The source of these requirements can be found in article 6 of the Nuremberg Code, which states that ‘The degree of risk to be taken should never exceed that determined by the humanitarian importance of the problem to be solved by the experiment’ (Code, 1949). The authors of the Belmont Report noted the difficulty of applying such an assessment:

Unlike ‘risk’, ‘benefit’ is not a term that expresses probabilities. Risk is properly contrasted with the probability of benefits, and benefits are properly contrasted with harms rather than risk of harms. Accordingly, so-called risk/benefit assessments are concerned with the probabilities and magnitudes of possible harms and anticipated benefits. (Commission, 1978, p. 8).

A standard definition of harm is not provided in the Common Rule, but the concept has been described extensively in the literature. For example, in clinical drug trials, the concept of harm is understood in physical, moral, and emotional terms in relation to subjects as a result of the trials (Shaw, S. & Barrett, 2006). In the social sciences, harms are typically either subjective experiences of distress, embarrassment and anxiety or non-subjective effects such as self-incrimination and reputational damage. The level of impact on an individual research subject may differ from person to person and is difficult to control for (Boddy et al., 2010). Scholars have also

suggested the following values as starting points from which to understand risk: physical injury, embarrassment, discomfort, inconvenience, stigmatisation, indignity, breaches of confidentiality, exposure to threats, self-esteem issues, privacy violations, insurability, and 'other psychosocial harms' (Guba & Lincoln, 1989; Kopelman, 2004).

'No More Than Minimal Risk'

The Common Rule enables an expedited review process in §46.110 if the study under review is deemed to pose 'no more than minimal risk'. When an IRB determines that the execution of a research proposal presents no greater than minimal risk to participants, it not only allows the informed consent to be waived according to §46.116(d)(1), but also allows research ethics reviews to be expedited and vulnerable populations to be included as participants in the study (such as children, prisoners, and pregnant and incapacitated persons; see the section on autonomy for a more detailed discussion). The Common Rule defines the concept of minimal risk in §46.102(i) as follows:

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

This standard of 'no more than minimal risk' has been called 'the most pivotal concept' in research ethics regulation (Kopelman, 2004). However, some academic debate exists regarding the inherently subjective interpretation of this concept, the

amount of discretion that an IRB has in weighing value-tensions, and where to set the threshold of risk.

The Common Rule allows two possible interpretations of the threshold of ‘minimal risk’ in §46.102(i): the risks encountered in ‘daily life’ or ‘during the performance of routine physical and psychological examinations or tests’. The first interpretation is vague and the second is directly related to biomedical or psychological practice, with little applicability outside of clinical practice.

An absolute interpretation of the ‘daily life’ risks standard would make the probability and magnitude calculation the same for all human subjects involved in research based on the experiences of a (non-existent) average person. This is, however, the accepted interpretation for US research (Tauer, 2002; Services, 2009). It assumes that all persons have the same standard of risk in everyday life, which is untrue. For example, research subjects who are at greater risk of harm would be treated the same as those with a reduced or ‘average’ risk of harm, which would not provide them with adequate mitigation of the risks that stem from the research. These issues are amplified when the population of research subjects is dispersed around the world in a wide variety of political environments and socioeconomic positions. In summary, the assessment of risks and benefits is not an absolute science, as it combines an unknown quantitative measure to a large range of qualitative categories.

Assessment

Scholars and research oversight committees are tasked with predicting or measuring the probability of harm occurring and anticipating its magnitude, weighing it against

the individual and social benefits of the research outcomes, and making a judgement based on these assessments. As described above, the types of harm may neither be quantifiable nor value-free (Shaw, S. & Barrett, 2006; Parker, & Barrett, 2003; Evans, 2000). It is therefore understood that an assessment of the benefits and risks of harm in a research proposal cannot be precise, and must be done on a best-effort and reasonable basis (Shaw, S. & Barrett, 2006). Furthermore, researchers and IRBs cannot guarantee participants' safety fully, so the role of oversight has become to minimise the potential for harm in the research (see the section on the US research ethics framework).

Research ethics oversight bodies are typically permitted some discretion in making their assessments (Levine, 1988; Shaw, S. & Barrett, 2006). The ideal level of discretion is itself also debated, however, raising concerns about different standards (Resnik, 2005) and conflicts of interests between the value of research and whether IRB members can be custodians of the public's rights in this regard (Sharav, 2003). In addition, when activities contribute to social and technical advances that are accompanied by unknown risks of harm, scholars have noted that institutions tasked with oversight become increasingly concerned with the management of risks (Beck, 1992; Power, 1997). Scholars have noted and complained that, as a result of complicated assessments, research ethics committees have become risk-averse and overly bureaucratic (for example, see (Deakin, Alexander, & Kerridge, 2009; Dingwall, 2006; Edwards, Ashcroft, R., & Kirchin, S., 2004; Minnis, 2004)).

Autonomy

The research ethics concepts of informed consent and vulnerable persons originate in the foundational value of autonomy. This section assesses the concept of autonomy and how it has been conceptualised (and criticised) in scholarly literature. The

following subsections discuss how autonomy can be operationalised by obtaining informed consent and approaches to conducting research with data subjects who may have diminished autonomy and are therefore vulnerable populations.

Autonomy in Society

The human subject model of research ethics is grounded in the concept of autonomy, and specifically in safeguarding the autonomy of research subjects. Indeed, autonomy is a fundamental value in the liberal tradition of modern political philosophy. The concept of autonomy underpins European and US principles such as the social contract and popular sovereignty, which are considered to enable a collective expression of rational choice whereby the institutions of political power are instrumental in the achievement of citizen welfare (for a further discussion of these principles, which are beyond the scope of this thesis, see (Mill, 2015; Rawls, 1972; Rawls, 2005)).

In the literature on moral philosophy and ethics, Immanuel Kant famously argued that humans are obliged to act towards others out of fundamental respect for their autonomy. His categorical imperative concept determines that individuals should be treated '[...] never merely as a means to an end, but always at the same time as an end' (Kant, 1785). This rule laid the foundations for deontology, one of the normative ethics reasoning methods. Rosamond Rhodes interpreted this rule to mean that there is no moral problem in using the services of professionals or research subjects to further one's own end if these persons are allowed to decide whether or not to participate (Rhodes, 2005). For Kant, autonomy was the key to understanding the moral obligations that people have towards each other.

In comparison, John Stuart Mill considered autonomy to be ‘one of the elements of well-being’, which is relevant for utilitarian calculations, another reasoning method in normative ethics (Mill, 2015). John Christman succinctly summarised this concept as follows:

Individual autonomy is an idea that is generally understood to refer to the capacity to be one’s own person, to live one’s life according to reasons and motives that are taken as one’s own and not the product of manipulative or distorting external forces. (Christman, 2008)³¹

Relevance for Research Ethics

Ruth Faden and Tom Beauchamp considered the pursuit of living up to one’s autonomy to be ‘a lofty ideal’ that cannot be achieved by everyone at all times (Faden, & Beauchamp, 1986). Still, they considered it to be a useful default position when engaging with other people, such as in research. Faden and Beauchamp therefore understood autonomy in the realm of research ethics as “restrict[ing] human subject research to those who (1) intentionally, (2) with understanding, and (3) without controlling influences agree to be subjects”.

Building on these philosophical ideas, the Belmont Report identifies autonomy to be a constitutive element of its principle of ‘respect for persons’, whereby it states that

Respect for persons incorporates at least two ethical convictions: first, that individuals should be treated as autonomous agents, and second, that persons

³¹ Online publication only, available at: <https://plato.stanford.edu/entries/autonomy-moral/>.

with diminished autonomy are entitled to protection [...]. (Commission, 1978, p. 4)

The Belmont Report further explains that autonomous persons are capable of deliberation about their personal goals. Acting in violation of someone's autonomy is to disregard another person's judgement. When a person is not in a position to make judgements and is thus considered to be vulnerable, the researcher has an obligation to protect that person based on a calculation of the benefits and risk of harm.

European research ethics guidelines contain similar statements to those in the Belmont Report. For example, the guidelines suggest that the primary way to ensure autonomy is to obtain the informed consent of research subjects (for example, see (Hughes et al., 2010)). These documents refer to vulnerable persons as being 'deprived' of their autonomy.

Informed Consent

Robert Veatch linked the underlying philosophy of the concept of autonomy to an actionable principle for research ethics when he stated that 'we should not treat people as means without their consent' (Veatch, 1998). Indeed, the Belmont Report and the European guidelines for research ethics both require the informed consent of human subjects prior to data collection. This is considered the mechanism by which to establish a respect for autonomy. By giving their informed consent, research participants are believed to indicate their choice to participate in a research project. This is considered to be meaningful when research subjects have been informed of

the potential risks, benefits, and alternative options and can withdraw their participation at any time (Anderson & Arsenault, 2005).

Connie Varnhagen et al. added that in a physical face-to-face meeting between researcher and subject, the researcher can assess whether a subject has made an informed decision, whether the subject has comprehended the information provided (possibly by answering the researcher's questions), and whether the subject's participation is indeed voluntary. Varnhagen et al. considered dialogue to be an important aspect of the informed consent procedure, because research subjects do not necessarily read the provided information (Varnhagen et al., 2005).

The Common Rule does not provide a definition of informed consent, but prescribes it in §46.116 as follows:

Except as provided elsewhere in this policy, no investigator may involve a human being as a subject in research covered by this policy unless the investigator has obtained the legally effective informed consent of the subject or the subject's legally authorized representative. [...]

This requirement follows from the principle of respect for persons as formulated in the Belmont Report, recognising human subjects as autonomous beings who willingly decide to be participants in research projects.

The informed consent process consists of three elements: information, comprehension, and voluntariness. While the central theme of this thesis is research in which no informed consent is sought, it is still useful to briefly describe these steps before the formal procedure to waive informed consent is described:

- Information: an expectation to disclose the research procedure, its purpose, anticipated benefits and identified risks, and the opportunity to ask questions and withdraw from the research at any time.
- Comprehension: giving the data subject the ability to understand the information and make an informed choice, by presenting it in an organised manner, allowing sufficient time for consideration, and allowing questions to be asked.
- Voluntariness: consent that is free from coercion (e.g., the threat of harm) and undue influence (e.g., offering an inappropriate or excessive reward) (Commission, 1978).

The 1949 Nuremberg Code stated that every research project required the informed consent of all of its subjects. This was the reaction of the presiding judges and their medical and scientific advisors to the involuntary and forced nature of experiments conducted on prisoners in Nazi concentration camps during the Second World War (Weindling, 2001). Scholars later challenged this absolute requirement, however, because this excluded many beneficial biomedical research projects in which it is not possible to ask subjects for their consent for a variety of reasons (especially when these projects involve children and vulnerable populations). The Declaration of Helsinki softened this absolute requirement. As a result, the Common Rule allows for the requirement of informed consent to be waived by an IRB under four specific circumstances, as described in §46.116(d):

- 1) The research involves no more than minimal risk to the subjects;

- 2) The waiver or alteration will not adversely affect the rights and welfare of the subjects;
- 3) The research could not practicably be carried out without the waiver or alteration; and
- 4) Whenever appropriate, the subjects will be provided with additional pertinent information after participation.

All four exemptions are relevant to this thesis. The first and second exemptions are related, since an adverse effect on the rights and welfare of subjects would pose more than minimal risk to them. These exemptions are discussed in more detail in the next subsection. The third and fourth exemptions for informed consent are mostly related to deception in research, whereby it is crucial to withhold information about an experiment or study from participants until after the experiment and data collection have been completed (Elms, Erwin, Gendin, & Kleiman, 1994). To qualify for these methods of waiving consent, researchers must explain why it is impractical to obtain informed consent, which is a stricter standard than the process of obtaining consent being merely difficult or inconvenient (Dittrich, Kenneally, & Bailey, 2012).

The Declaration of Helsinki allowed researchers to use proxies or representative authorities to replace individual informed consent procedures.³² Some (vulnerable) groups may lack the ability or opportunity to give their informed consent, so their interests may be safeguarded by those who are responsible for them. The Common Rule limits such a proxy to a 'legally authorized representative', which is defined as 'an individual or judicial or other body authorized under applicable law to consent

³² See paragraphs 27 – 30 of the Declaration of Helsinki, is available at: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>.

on behalf of a prospective subject to the subject's participation in the procedure(s) involved in the research.'

Some scholars have questioned the extent to which an informed consent procedure can, in fact, operationalise autonomy meaningfully. While a useful moral insight, Rhodes argued that constituting the atrocities of the Nazi concentration camps merely on a lack of informed consent does not do justice to the conditions under which prisoners in concentration camps were coerced or even forced into experimentation. She warned that the idea of informed consent by itself being the mechanism by which to establish or respect autonomy places too much responsibility on the concept (Rhodes, 2005). Furthermore, several studies have demonstrated that research subjects typically skim-read the documentation, if at all, and may be unable to recall or even comprehend the information provided in the informed consent procedures (Varnhagen et al., 2005).

Vulnerable Persons

The concept of vulnerable persons is central to this thesis, because unsuspecting internet users (as subjects in an experiment) may be considered as having diminished autonomy. Research ethics frameworks and guidelines typically require special precautions when a study includes vulnerable people, although the scope and meaning of the concept vary (Bracken-Roche, Bell, Macdonald, & Racine, 2017). While there is widespread agreement that vulnerable populations should receive special consideration and protection in research, the codification of vulnerability into research ethics frameworks has been described as too vague and, at the same time, too broad (Schroeder, & Gefenas, 2009). This is unsurprising, because the concept of vulnerability can be understood in a variety of ways. For example, Martha Albertson

Fineman considered vulnerability to be a general and 'universal, inevitable, enduring aspect of the human condition' (Fineman, 2010). She described how all persons are vulnerable to some extent within the institutions developed by societies. In his philosophical work entitled 'Protecting the Vulnerable', Robert E. Goodin considered the concept of vulnerability in terms of power balances, the potential to exploit or inflict harm, and the resulting responsibilities (Goodin, 1986). Samia Hurst interpreted the concept of vulnerability in research ethics, specifically in health care, as 'an identifiably increased likelihood of incurring additional or greater wrong' (Hurst, 2008).

Based on considerations like these, legislators prescribed particular caution for studies that use vulnerable populations. To be effective, research ethics frameworks must offer operationalisable definitions and instructions to researchers and ethics review boards. Thus, vulnerable people have been defined as 'clearly lacking decisional capacity' (Rhodes, 2005) and being 'deprived' of their autonomy (Hughes et al., 2010; Europe, 2012). For example, the US research ethics framework categorised specifically 'children, prisoners, pregnant women, mentally disabled persons, or economically or educationally disadvantaged persons [...]' as vulnerable. In the Code of Federal Regulations for the Protection of Human Research Subjects, §46.111(b) prescribes that when a study includes vulnerable persons, the IRB should assess whether "additional safeguards have been included in the study to protect the rights and welfare of these subjects".

A European research ethics guideline document provides the following three elements to identify a research subject as being vulnerable and therefore in need of additional care, which again alludes to a lack of autonomy:

- 1) Subjects who lack competence will be unable to protect their interests by choosing to give or withhold consent [...].
- 2) If the voluntariness of the subjects' consent is compromised, this may similarly prevent them from choosing to give or withhold consent in a way that would protect their interests.
- 3) The physical (or psychological) condition of some subjects leaves them especially liable to harm, for example as a result of frailty through age, disability, or illness (Hughes et al., 2010).

Some academic debate exists on the chosen definitions in research ethics frameworks. Dearbhail Bracken-Roche et al. found that a conceptualisation and operationalisation of the concept of vulnerability is lacking, which had led to misguided discussions about policy refinements in research ethics. Specifically, they argued that interpretations that are too narrow (such as the list of groups provided in US law) exclude some groups that may require special protections, whereas a definition that is too broad creates confusion about how to apply 'special protection' and to whom (Bracken-Roche et al., 2017).

Rosamund Rhodes put further pressure on the current formulation of protections for vulnerable persons, arguing that autonomy is not respected when protecting groups that are not considered to have autonomy in the first place. Instead, she argued that protecting vulnerable groups amounts to paternalism rather than a respect for autonomy. Rhodes claimed that creating blanket special protections for vulnerable groups denies them the opportunity to consider the benefits and costs from their particular perspectives (Rhodes, 2005). To support this criticism, Florencia Luna and

Sheryl Vanderpoel noted that the state of being vulnerable is relational and dynamic, depending on a person's context and the situation. Indeed, vulnerability should be understood in different social and contextual layers rather than as a binary checklist (Luna, 2009; Luna, F. & Vanderpoel, 2013).

Deception

Deception research occurs when a researcher intentionally misinforms participants about the nature of a study or does not (fully) disclose relevant information when traditionally informed consent would be required (Elms, Erwin, Gendin, & Kleiman, 1994). The underlying assumption is that the relative benefits of conducting the research with misinformed research subjects will outweigh the costs of participants being temporarily deceived (Pittenger, 2003). Guidelines for deception research typically require the following to be true:

- 1) The research provides considerable scientific, educational, or applied value;
- 2) No alternative methods exist to achieve the expected results;
- 3) Risks to deceived participants are minimal;
- 4) Deception does not target significant aspects that would affect the willingness of potential data subjects to participate, such as physical risks, discomforts, or unpleasant emotional aspects;
- 5) Participants are debriefed after the research and are asked for their consent to use their data in the study (see (APA, 2002; Elgesem, 2002)).

It is believed that the potential harms of deception research, such as discomfort or negative emotional reactions, can be rectified by a debriefing after the data has been collected from the deceived research subject (Patterson, 2010). David Holmes suggested that deceived research subjects should be 'dehoaxed' and 'desensitised'. By 'dehoax', Holmes means that participants should be informed about the deceptive element of the study and that the rationale for employing such a methodology should be explained to them. If any negative feelings or emotions may have arisen, the researcher can 'desensitise' the research subject by discussing the distress (Holmes, 1976). Joan Sieber adds that the debriefing should provide participants with a feeling that they have contributed to society and scientific understanding by participating in the deceptive study (Sieber, 1983, p. 1).

Diana Baumrind considered deceptive research practices to be intrinsically in violation of research ethics, specifically in her discipline of psychology. While she saw value in the debriefing considerations suggested by Sieber, she argued that deception research in which fully informed consent is waived infringes the obligation of researchers to protect the welfare of the research subject and the obligation of researchers to be trustworthy (Baumrind, 1985; Baumrind, 2015). She also argued that the harm that deception does to research subjects, the research (and psychological) profession, and society does not justify the incremental gains in scientific knowledge.

Privacy

The concepts of publicly available data, confidentiality, and anonymity are part of the umbrella of the broader concept of privacy. This section and its sub-sections only provide a provisional discussion of the concept of privacy as it relates to research ethics. The case studies return to the definitions and concepts described in this and the next chapter. These concepts are interpreted, tested, and potentially strengthened or reinterpreted in this thesis.

Privacy and confidentiality have been the cornerstones of the patient-physician relationship since the introduction of the Hippocratic Oath around the 3rd century BC (Moskop, Marco, C. A., Larkin, G. L., Geiderman, J. M., & Derse, A. R., 2005). More recently, these concepts have gained in importance in the subject-researcher relationship. The Belmont Report, for example, has been interpreted to individual privacy and autonomy are necessary to honour the three key principles of the report: respect for persons, beneficence, and justice.³³

The right to privacy is also encoded in several important pieces of legislation in the international human rights framework, as well as taken up in national constitutions around the world. Principle 9 of the Declaration of Helsinki states the obligation for researchers to guarantee the privacy and confidentiality of research subjects' personal information. Similarly, the Ethical Guidelines for Biomedical Research Involving Human Subjects provided by the Council for International Organizations of Medical Sciences (CIOMS) and the World Health Organization (WHO) stipulate that researchers have similar obligations (WHO, 2002).

³³ For example, see the University of California's Office of Research interpretation, available at: <https://www.research.uci.edu/compliance/human-research-protections/researchers/privacy-and-confidentiality.html>.

The section of this thesis on the US research ethics framework discusses how the US research ethics Common Rule (Code of Federal Regulations Title 45 Part 46) regulates the use of participants' private information. The Common Rule states in §46.111(7) that IRBs assess that, 'When appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of data'. European researchers are primarily bound by the General Data Protection Regulation, which provides strict rules for the collection and use of individuals' personal data.

Despite the extensive scholarly literature on the concept of privacy, it remains a vague concept which has been found to be difficult to define precisely (Post, 2000; Solove, 2006; Thomson, 1975). Lincoln and Guba found that concepts of privacy have been too narrowly defined in research ethics, leaving many private spaces of persons open to investigation by researchers (Guba & Lincoln, 1989). Daniel Solove attempted to give some guidance in relation to the meaning of privacy in *A Taxonomy of Privacy*, but even he stated that 'Privacy is a concept in disarray. Nobody can articulate what it means' (Solove, 2006). The concept of privacy may thus be better understood as an umbrella term, referring to a group of related concepts, issues, and values. Privacy has been subdivided into several "separate but related concepts": information privacy, bodily privacy, privacy of communications, and territorial privacy (EPIC, 2007).

While several types of privacy may be violated by researchers, research ethics frameworks typically focus on the collection, use, and reuse of data. Academic disciplines may uphold further specific ethical standards, such as the respect for

physical integrity, if this is relevant for their experimentation, but information privacy remains the most applicable for research in general.

The concept of information privacy is understood to be a means to control personal data (Altman, 1975; Cohen, 2000; Westin & Ruebhausen, 1967), to set personal boundaries as a form of autonomy (Altman, 1975; Cohen, 2012), to use personal data as a commodity (Garfinkel, 2000; Laudon, 1997; Van Dijk, 2010), or as a social value (Regan, 1995; Solove, 2016), which is highly contextual in nature (Acquisti & Grossklags, 2004; Altman, 1975; Bennett, 1991; Margulis, 1977; Margulis, 2003; Solove, 2004; Solove, Rotenberg, & Schwartz, 2006; Westin & Ruebhausen, 1967). In summary, information privacy is itself an unclear concept.

The need for information privacy may be better understood by describing the harms that may occur when it is violated. For example, sensitive research data in the possession of certain parties—such as identity thieves, malevolent (possibly authoritarian) governments, abusive spouses, aggressive marketers, etc.—may lead to serious financial, reputational, or physical harm, such as exploitation, embarrassment, or even punishment (Frankel, M. S. & Siang, 1999). The sensitivity of data and information differs per context, whereby seemingly innocuous data or inferences in one context can lead to imprisonment or social exclusion in another (Thomas, 1995). Furthermore, Maurice Punch stated that some of these harms caused by privacy violations may only materialise several decades after the data has been collected (Punch, 1994), which is difficult to assess or manage in a research ethics review procedure.

Existing and Publicly Available Data

The Common Rule offers several exemptions to research ethics review in §46.101(b), of which the most relevant concern the reuse of existing and publicly available information in sub-part (b):

Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

This provision exempts research that is specifically based on the reuse of existing data or that uses information from publicly available sources. At the time of drafting the Common Rule, lawmakers could not have foreseen the extensive capabilities of data analytics and resulting re-identification of individuals in seemingly anonymised datasets. Before the advent of digital technologies and networks, it was relatively obvious which information would be publicly accessible; the average person could either access it or not.

Confidentiality

Confidentiality is understood to be the mechanism by which research subjects' data is protected by the researcher. While privacy relates to the extent of data collection by the researcher, confidentiality obligations regulate the extent to which the researcher may disclose the collected data.³⁴ In the research context, this has become

³⁴ Online publication by the University of Columbia, available at: <https://cenmtl.columbia.edu/projects/cire/pac/foundation/>.

understood to mean that researchers may not discuss the information provided by a participant with others, and that research findings must be presented in a way that ensures individuals may not be identified (Wiles, Crow, Heath, & Charles, 2008).

Confidentiality requirements in research ethics are based on the principle of respect for persons. Some scholars have argued that the researcher's obligation to protect confidentiality is, in fact, a higher social duty than the physician's or clinician's duty to protect confidentiality, due to the voluntary nature of research participation.³⁵

Lowrance argued that only in very pressing situations may research data be disclosed, such as in a public health investigation or crisis about which a researcher may possess vital information (Lowrance, 2003).

Anonymity

When data subjects are individually identifiable, research ethics frameworks and data protection laws require that the data collection process is scrutinised, and the full extent of the researcher's responsibilities and obligations may become applicable. Such regulatory frameworks do not apply, however, when the data subjects are rendered anonymous in a way that the human to whom the data relates is not identifiable. The concept of confidentiality is closely connected to the concept of anonymity, although the anonymisation of datasets does not necessarily address all the issues that relate to confidentiality. The re-identification of datasets has become a more pertinent issue in internet research, which is described in the conceptual framework of this thesis.

³⁵ Id.

Part 4: Law, Ethics, and Research Themes

This part of the chapter lays the groundwork for the socio-legal thesis that takes into account underlying ethical values and frameworks. Research ethics law—as the name implies—is a legal framework that operationalises ethical theories and values with regards to research practice. The US Common Rule is an example of such a law, which codifies the responsibilities and moral tensions faced by researchers and prescribes a procedure for the review of research. To lay the groundwork for this thesis, this part is itself divided into three distinct sections.

The first section describes the development of law that is explicitly underpinned by ethical frameworks, or law that intends to promote particular values. Some scholars note that this form of law is novel, responding to the needs of modern societies that are based on value pluralism and fast moving innovations. However, law and ethics have always been closely intertwined and their distinctions are not clear-cut. This section does not intend to describe the concept of law or its relationship with ethics in detail. Instead, the section merely positions the Common Rule within a wider array of legal frameworks, and tees-up the next section on ethical theories, which provides the lens through which the case studies of the thesis are assessed.

The second section explores the concept of ethics and its various theories and frameworks that are relevant to research ethics and this thesis. The section explores how the various ethical theories are relevant for the law, and differentiates between those that are most relevant to legislating and those that are most relevant to interpreting the law in a particular situation. The descriptions of the ethical theories are limited to a high-level overview and do not engage with the multitude of interpretations and disagreement in academic ethics.

The third section describes the main ethical theory used by this thesis—the ethics of care—in more detail. The section describes its relatively short history in some detail, before it divides up its core tenets into the research themes that will be used throughout the thesis.

Law

The codification of ethical theories and principles in legal frameworks, such as the Common Rule, requires a brief introduction. Law and ethics are separate concepts, although mutually reinforcing. Their relationship is closely aligned to the differences between justice and morality. To identify and articulate the differences, though, depends on an interpretation of several competing legal and ethical theories (Bondolfi, 2000). For example, natural law theory and legal moralism assert that law is an expression of morality. Legal positivism, on the other hand, contends that law is a command from a sovereign and that there is not necessarily a connection between morality and enacted laws. This thesis does not attempt to take a position in those debates. It is, however, useful to distinguish between a variety of legal and ethical approaches, to pinpoint exactly where this thesis is contributing.

Traditional role of law

Many of the concepts, themes, and norms in this thesis are elusive and probably impossible to define precisely. The same holds true for the concept of law (Arnold, 1937). This section provides a very brief overview of the evolving concept of law. The next section describes more modern conceptions of law, that explicitly encode ethical theories and attempt to impose values on stakeholders.

In short, a traditional view of law considers legislation as a social construct that is developed by (democratically) legitimate social institutions and is enforced by authorities with the aim to regulate the behaviour of their subjects (Galligan, 2006). John Austin put forward a 'command theory' of law in 1832, in which he suggested that laws are the commands of a sovereign that are intended to control behaviour and thus society (Austin, 1863). In 'The Concept of Law', H.L.A. Hart added a democratic twist to this theory, suggesting that primary (obligations) and secondary (procedural) laws exist, whereby the secondary laws legitimise the procedures to establish and enforce the primary laws (Hart & Green, 2012). Leslie Green further interprets Hart's work by stating that the primary obligations in law often represent moral norms or principles, which, when codified, provide institutional support, remedies, and penalties when they are not lived up to (Green, 2013). Prof. Tom Tyler argues that the legitimacy of these institutions is based on their subjects' perceptions of them, which in turn are based on the subjects' judgements about the fairness of the procedures used by the institutions and authorities (Tyler, 2006).

In this traditional paradigm, the dominant moral norms in a particular context were considered to be clear and assigned legal status by law. Laws typically operate on the assumption that its subjects are reasonably competent and independent people who want their rights to self-determination and autonomy protected. Jonathan Herring analysed that "[t]hese are seen to promote an individualist conception of people" (Herring, 2017, p. 165).

This approach and assumption has been increasingly criticised, accusing law from losing touch with the lived reality of its subjects. Carol Smart states that "the rights approach takes and translates personal and private matters into legal language. In so

doing, it reformulates them into issues relevant to law rather than to the lives of ordinary people" (Smart, 2000, p. 237). Society and its legislative bodies are currently being confronted with new innovations at a rapid pace. Furthermore, scholars have argued that indisputable moral norms that give guidance no longer exist in modern societies, and the law does not appear to be as effective an instrument to regulate behaviour as it may have been in the past (Van der Burg & Brom, 2000; Held, 2005)}. Policy vacuums have appeared in areas such as biotechnology, the Internet, health care, and animal rights, where legislators typically consult ethicists to provide guidance. It is in these fields that the moral dimensions of legislative review or legislative intervention are relevant.

Expressive and communicative function of law

Van der Burg and Brom note that some laws have become expressive and communicative of ethical theories and values as it adapted to modern challenges (Van der Burg & Brom, 2000). First, they describe laws as expressing fundamental standards of norms and values. These are political statements of the values to which a particular (group within) society aims to adhere, but it leaves citizens to determine how to interpret those values in concrete situations. Second, the law communicates a procedural dimension with which it creates and supports the institutions that promote the implementation or further discussion of how to interpret, elaborate, and implement values.

Following the idea of law as an expressive and communicative medium for theories and values, Jonathan Herring describes how law can be "used powerfully to encourage people to act in a way that will promote caring relationships", instead of promoting only self-determination and autonomy as virtues. In his paper, he describes that it may not be possible to devise enforceable laws that prescribe that a

nurse must act in a caring way towards patients. However, law can create the "environment and processes that might facilitate compassionate caring relations", just as "contract law plays a role in creating an environment in which contracts are respected" (Herring, 2017, p. 164).

As a result of these new functions, new governance phenomena have appeared. These include ethics committees, ethical codes, ombudsmen, covenants between industry and governments, international conventions, and international organisations. The Common Rule and the ethics review boards (IRBs) it has instituted are an example of the expressive and communicative role that law has taken. Their procedures and requirements place responsibilities and obligations on the relatively powerful researchers in their relationship with more vulnerable research subjects.

Indeed, the relative success of the Common Rule shows that it is possible to promote ethical theories and values through law. However, this thesis aims to uncover whether these underlying theories and values are still appropriate in the context of research ethics laws applied to the review of technical Internet research. The next section describes some common ethical theories and values.

Ethics Frameworks

Ethics as a discipline is concerned with understanding right from wrong. It is synonymous with the sub-branch of philosophy called moral philosophy, which reasons about morally right and morally wrong actions. James Fieser defines ethics as '[...] systematizing, defending, and recommending concepts of right and wrong behavior' (Fieser, 2018). It is a rigorous discipline with deeply-established theories based on several millennia of recorded thought.

Ethics consists of a number of study areas, ranging from abstract to concrete. It is beyond the scope of this thesis to interrogate all the constituent parts of ethics as a discipline. It suffices to distinguish between the following four categories of ethics:

- **Metaethics** is concerned with the nature of ethical concepts, such as ‘right’ and ‘wrong’ (Baggini & Fosl, 2007). It is the abstract study of the meaning of words, concepts, statements, and arguments used in ethics (Van Roojen, 2015).
- **Normative ethics** is concerned with the reasoning processes about the particular concepts, actions, and things that may be right or wrong (Baggini & Fosl, 2007), and seeks to set norms or standards for conduct (Singer, 1985). A variety of moral reasoning frameworks have been developed over the centuries, such as utilitarianism, deontology, virtue ethics, and care ethics (for example, see (Anscombe, 1958)).
- **Applied (or practical) ethics** applies normative frameworks to real-world actions and situations. It is the study of normative theories and the application of these sets of principles to real-world practical moral problems (Singer, 1985). Practical ethics typically underpins further sub-disciplines, such as medical ethics or bioethics (Singer, 2011).
- **Professional ethics**—or ethics compliance frameworks—command specific procedures and behaviours (Treviño, Weaver, Gibson, & Toffler, 1999). These frameworks exist for professions (e.g., medicine, law, and research) or for sectors (e.g., government, finance, and research).

Normative Ethical Frameworks

Research ethics cuts through the variety of the categories above. A research ethics law may be a version of professional ethics, though its rules are underpinned by normative ethics and its moral reasoning theories. This section discusses the normative frameworks that have been relevant in the literature of research ethics law. These are consequentialism, deontology, virtue ethics, and the ethics of care. Other frameworks include (but are not limited to) contractualism, hedonism, naturalism, and pragmatism.

The four main frameworks are described briefly below and in two categories. Consequentialism and deontology are discussed as *abstract* and *universal* frameworks. Virtue ethics and the ethics of care are discussed as *contextual* and *relational* frameworks. The section ends with an explanation that the former has been considered to be more useful for legislating, while the latter may be more useful for interpreting appropriate responsibilities and duty in a particular situation.

Abstract and Universal Frameworks

Consequentialism refers to ethical frameworks that judge the moral nature of an action entirely by its consequences. The basic claim is that the consequences of our actions are crucial in determining the morality of the actions. The intentions or motives of the agent are not directly relevant to the moral nature of the actions. The most prominent interpretation of consequentialism is called *utilitarianism*, which focuses on maximising general pleasure (or happiness). Jeremy Bentham described utilitarianism as ‘the sum of all pleasure that results from an action, minus the suffering of anyone involved in the action’ (Bentham, 1789, #29161). This implies a moral calculation whereby every action is judged to be right or wrong, based on how it affects the net balance of pleasure over pain (or happiness over unhappiness) in

the universe. Similar utilitarian frameworks substitute happiness for the greatest welfare, or whether the action allows the greatest number of people to live according to their preferences.

Consequentialism and utilitarianism are criticised for sacrifices of detail in its abstract moral accounting, their disregard for individual rights, and the likely inaccuracy of moral calculations. Optimising for the greatest amount of happiness, pleasure, welfare, and preferences for the greatest number of people invariably means that some people's quality of life is sacrificed or diminished. The ends thereby justify the means. Further, it is difficult—if not impossible—to make consequentialist moral calculations which take into account the subjects' benefits and harms in terms of their social, economic, and political effects. Critics further assert that it is too difficult to fully grasp consequences ahead of an action, especially in complex societies (Scheffler, 1988; Scheffler, 1994; Smart & Williams, 1973).

Deontology is a set of ethical frameworks that take duty and rights as a starting point in moral reasoning. The source of duties (or rules) may be internal to the social situation or externally imposed by laws or other sources of power, such as religious doctrine or cultural norms. Deontological ethics frameworks determine that an action should be judged according to whether the action itself is right or wrong under particular rules, irrespective of its consequences (Singer, 1985). Deontological frameworks are therefore more concerned with the justification of the means, and less with the ends (or consequences). The dominant deontological framework is Immanuel Kant's categorical imperative. In his body of work, Kant argues that actions must not be influenced by desires or the identity of the person making a moral deliberation. Furthermore, actions must be compared to whether they would

be desirable if they were to become a universal law. In addition, Kant argued that individuals should be treated '[...] never merely as a means to an end, but always at the same time as an end' (Kant, 1785, #99762).

A common criticism of the deontological adherence to rights and duties is that the frameworks do not offer a clear path to a solution when values clash (Miller, 2010; Sedgwick, 2012). Furthermore, critics have found the disregard for consequences to be troublesome, because blindly complying with duties may result in harmful consequences for some (McMahon, 1991). In response, deontologists have argued that in particular situations, when values clash, one duty can have a greater claim on a person than another duty, especially if the person can avoid harmful consequences (Ross, 2011).

Both consequentialism and deontology offer reasoning frameworks are based on assumptions that an agent is independent, rational, seeks a fair outcome, and that the solutions are universally applicable. Applying these frameworks demands an abstract view of society and hypothetical contexts. Derek Parfit argued that relentless ethical efficiency in decision-making based on these frameworks would defeat some of the most precious aspects of the human experience, such as maintaining close personal relationships and building character (Parfit, 1984). The complexities of moral calculations or adherence to specific duties are unlikely to serve agents well when confronted with concrete, real-life problems. Therefore, the abstract and universal frameworks are useful tools to support the creation of high-level policies and legislation, but not necessarily for making decisions in particular situations or contexts.

Contextual Frameworks

Virtue ethics refers to a range of ethical theories in which the concept of virtues is treated as a fundamental quality of character. These frameworks focus on the character of the actor, or the *virtuous agent*. They prescribe “how we should be rather than what we should do” (Darwall, 2005, p. 1). The source of morality is therefore found *within* the moral agent.

The most dominant virtue ethics framework remains the one offered by Aristotle during his lifetime (384–322 B.C.E.), which has influenced many virtue ethicists over the millennia. In his framework, moral virtues are understood to be elements of a person’s character, such as patience, moderation, courage, and honesty. Virtues thus act as a personal hermeneutic tool that helps individuals to interpret, define, and pursue the right action in particular situations, based on their own character, knowledge, and experience (Edgar, 2005, #56496). When confronted with a challenge, the “virtue directs, modulates, and integrates the enactment of a person’s individual moral virtues, adjusting their habitual expression to the unique moral demands of each situation” (Vallor, 2016, p. 20). People are not born with virtues but must cultivate them through experience and example.

The *ethics of care* emerged in the last few decades as an alternative moral theory to those described above. At its core, the ethics of care operates from the assumption that there is moral significance in human relationships, their dependencies, and vulnerability with regard to each other and institutions in society (Tronto, 1993). Moral responsibility is derived from the accidental or purposeful embeddedness in particular social contexts, the power dynamics within these situations, and the

vulnerability of those for whom the moral agent has a responsibility to care (Held, 2005). Consequently, moral harm derives from failures to respond to needs and taking appropriate responsibility, instead of merely violating rights or causing generally harmful consequences (Kittay, 2011).

As a developing moral theory, scholars are in the process of situating the theory and responding to criticism. For example, some question whether the ethics of care is not simply a virtue ethics framework which is limited to intimate settings of proximate human beings (Thomas, 2011). Scholars engaged in developing the ethics of care respond that, instead of a focus on the self (as with virtue ethics), the nature and quality of caring attitudes in hierarchical contexts of dependence are the focal point of moral analysis (Held, 2005). Further, the theory has been successfully applied to many contexts where actors are distant from each other and do not have a personal relationship, such as the welfare state (Engster, 2015) and global security policy (Robinson, 2011).

Both virtue ethics and the ethics of care are concerned with guiding action in specific contexts and situations. In contrast to abstract and universal frameworks, they explicitly include considerations of individuals and their relationships. They also include values such as emotion, empathy, sensitivity, wisdom, and courage. Therefore, contextual and relational ethics frameworks are better positioned to serve the contextual interpretation of laws, responsibilities, and duties that have been devised using abstract and universal and abstract ethical frameworks.

Virtue ethics and the ethics of care are useful lenses to assess the responsibilities and duties of Internet researchers. The focus on relationships as a source of morality

makes the ethics of care particularly suited for an analysis of the responsibilities of researchers. After all, this part of the chapter opened with a description of the various relationships an academic researcher has: to their data subjects (who may be implicated without their knowledge or consent), to society, and to their colleagues in the research profession. The next section reviews the ethics of care as a moral theory and develops the research themes that are used throughout this thesis.

Ethics of care: Background

The ethics of care is a relatively novel ethical framework that has been described as an “ongoing collective research project” (Bubeck, 1995, p. 10). Carol Gilligan's criticism of the use of ethics theories in developmental psychology with regard to morality is widely regarded as the starting point for the ethics of care as a moral theory. In her book ‘In a Different Voice: Women's Conceptions of Self and of Morality’, Gilligan laid the foundations of what would become the ethics of care (Gilligan, 1977). Nel Noddings built on Gilligan's work, describing the moral nature of the attitudes of those in the caring professions as being responsive to the needs of those cared for, and understanding their points of view (Noddings, 1982).

Over the years, the ethics of care literature departed from its feminist roots and developed into its own, distinct normative ethical framework. Scholars are increasingly referring to it as a “full-bodied moral and political theory” (Engster & Hamington, 2015). Collins clears up some misunderstandings of the un-gendering of the ethics of care when she states:

“To describe care ethics as the feminist moral theory is to present only one understanding of the theory behind feminism, and to present only one

understanding of the proper political and social uses of care ethics" (Collins, 2015, p. 7).

She aspires to develop the ethics of care "as a moral theory that applies to everyone, all of the time" (p. 8). In this thesis, I follow scholars such as Engster, Hamington, and Collins in applying the learnings of this novel moral theory to research ethics in a similarly novel technological environment. While I do not mean to dismiss the historical roots of this moral theory, gender differences or stereotypes do not play a direct role in the topic of this thesis.

The ethics of care contrasts itself from other normative theories by its focus on principles that are traditionally neglected or marginalised in other ethics theories. It starts from the view that an individual (or 'the self') is inherently relational, by which one's identity and autonomy are also relational and dependent on others. Moral obligations emerge from these relationships and dependencies. Collins boils the theory down to its essence when she defends that the central slogan of the ethics of care is (or should be) "dependency relationships generate responsibilities" (Collins, 2015, p. 2).

In contrast to other theories, the ethics of care does not provide a value maximising injunction for neutral agents (such as consequentialism). Neither does it constrain actions by looking for vague notions of rights as guidance (such as deontology). And further, the ethics of care derives responsibilities from existing relationships, not an ideal-type character stereotype (such as virtue ethics). This is not to claim that the ethics of care is superior, but merely that it is different and serves a unique purpose in moral theorising.

Some of the core tenants of the ethics of care have been articulated in Engster and Hamington's introduction to 'Care Ethics and Political Theory' (Engster & Hamington, 2015) and Stephanie Collins' 'The Core of Care Ethics' (Collins, 2015). Taken together these are (1) a relational approach to morality; (2) contextually-oriented deliberation; (3) responsiveness; and (4) the use of emotions and attitudes as moral tools. Together, these tenets form the research themes of this thesis. The concepts of the Common Rule and the issues arising in the case studies are assessed according to these themes.

Ethics of Care: Research Themes

Research theme 1: Relational approach to morality

As a starting point, the ethics of care literature views humans as fundamentally interdependent and relational (Engster & Hamington, 2015; Held, 2005). Our relationships are a "fundamental feature of our existence", which is so far underrepresented or even dismissed as irrelevant in other moral frameworks (Robinson, 1997, p. 4). The importance of a relationship is both the value placed on them by participants, and the consequences of the relationship for them. A beneficial relationship may be characterised as a positive, whereas an exploitative or oppressive relationship may be characterised as having negative value for (one of the) participants. Collins represents other scholars writing about the ethics of care when she suggests that relationships should be a source of moral deliberation:

"[...] relationships ought to be (a) treated as moral paradigms, (b) valued, preserved, or promoted (as appropriate to the circumstance at hand) and (c) acknowledged as giving rise to weighty duties" (Collins, 2015, p. 10)

Relationships come in many shapes and sizes, and do not necessarily only occur between equals who have volunteered to be in a relationship (Brannelly, 2018). In the ethics of care literature, a relationship is typically described as being towards an 'other.' The initial publications concerned mostly relationships between persons who had some form of *relatedness* and *proximity*. They were thus *particular* and *proximate* others (Held, 2005). Later scholars have questioned this assumption and argued that care relationships can exist between *distant*, *particular others* or even with a *generalised* and *distant* group who the carer may never meet in person (Bubeck, 1995; Held, 2005; Slote, 2007; Barnes & Brannelly, 2015). According to Slote and Held, the relatedness requirement is increasingly irrelevant in a society where the distant other is reachable or observable through technology, and actions in one place may have significant effects elsewhere. In fact, moral obligations to distant others (whether generalised or particular) have gained relevance in recent philosophical publications. Peter Singer, for example, argued that mere spatial distance should be irrelevant for moral theorising (specifically in the realm of humanitarian aid). In his essay 'Famine, Affluence, and Morality', he states:

"It makes no moral difference whether the person I can help is a neighbor's child ten yards away from me or a Bengali whose name I shall never know, ten thousand miles away. [...] The moral point of view requires us to look beyond the interests of our own society." (Singer, 1972, p. 231)

The role of the ethics of care towards distant persons is described in Fiona Robinson's work on interpreting the ethics of care in the global security domain. Her focus is on "those who are constructed as weak, vulnerable, and dependent". In her own words, her work "recognizes the complex interdependence and relationality that characterize relations among states, institutions and individuals even in distant

geopolitical regions.” (Robinson, 2011, p. 137). Other scholars have described the potential of the ethics of care to influence and shape social policy, which is similarly focused on distant and generalised others (Held, 2006; Noddings, 2002; Ruddick, 1995; Tronto, 1993). Michael Slote differentiates this “general humanitarian caring or concern about people one only knows about” to the other “intense personal caring towards people one knows” (Slote, 1999, p. 26). This positioning of the ethics of care is translatable to the context of technical Internet research ethics, where relatively powerful scholars conduct experiments in the name of institutions on relatively vulnerable data subjects that may be spread across the world. This approach to the ethics of care is concerned with informing the policies and institutional contexts necessary that enable caring relationships.

For this thesis, I understand the research theme ‘relational approach to morality’ to apply to the nature of relationships in which a researcher is inherently embedded.

These are:

1. The data collection relationship between the researcher and their **data subjects**, understood as distant, particular or generalised others.
2. The patronage and trust relationship between the researcher and wider **society**, embodied in part by their political representatives.
3. The collegial relationship between researchers and their **colleagues** in their institutions, their research discipline, and the research profession as a whole.

Research theme 2: Contextually-oriented deliberation

The ethics of care recognises that people do not live in isolation, but are inherently connected, relational, and dependent. The nature of a relationship is therefore characterised by the contexts within which it is situated. Relationships are shaped by

the participating individuals, and the existing historical, social, and political contexts in which they are nested (Minow & Shanley, 1996). An individual context can be characterised in a variety of ways. For example, a relative power imbalance, leading to vulnerability or dependence, is characterised as a hierarchical context (Sevenhuijsen, 2003). It would be oppressive if the relatively powerful person subjects the other to actions against their will. The broad set of relationships and contexts within which people are embedded give rise to nested dependencies and nested obligations (Kittay, 2001).

To understand the various contexts that need to be taken into account, it is useful to approach the characterisation of a relationship through the lens of lived experiences of individuals and their interactions or dependencies. An understanding of participants' experience, "including their history, relative power, relationships, and so forth" (Engster & Hamington, 2015, p. 5), and "understanding the lived experiences of those involved in a relationship" is required to assess the nature of a context (Brannelly, 2018). This allows a more truthful and informed understanding of moral obligations of persons within a relationship (Held, 2005). However, the other does not necessarily need to be a particular individual, because social positions and experience of those recognised as weak, vulnerable, and dependent can be generalised (Bubeck, 1995; Held, 2005; Brannelly, 2018)

For this thesis, I understand the research theme 'contextually-oriented deliberation' within two separate, but intersecting contexts:

- The first context is the research activity within which a researcher forms a relationship with an unsuspecting Internet user. The research context is characterised by the roles and relative power of the (unwitting) participants.

For example, a research context may be characterised as hierarchical and even as oppressive.

- The second context is the social and political environment within which the implicated Internet users are embedded. There are a wide range of variables that characterise their contexts, and their relevance depends on the nature of the research.

Research theme 3: Responsiveness

In the ethics of care literature, a moral harm is considered as the consequences that result from a failure of responsiveness (Kittay, 2011). Responsive practices involve both thought and action, which are interrelated towards an end (Tronto, 1993).

Responsiveness is characterised by identifying another's needs, understanding the need from their point of view, and interpreting a required action in the context of the relationship.

The starting point for responsive and caring action begins with a willingness to identify the needs of others and an intention to take responsibility for meeting their needs (Sevenhuijsen, 2003; Engster & Hamington, 2015). First, an agent must recognise the need and make a determination that the needs must be met (Tronto, 1993). Second, it requires that one understand a situation, take the point of view of the other, and listen to the moral deliberations of the other (Held, 2005; Sevenhuijsen, 2003). Third, Bubeck describes responsiveness as an attitude of "yielding flexibility" and a "capacity to be ready for surprise" and "resourceful interpretation" in responding to the needs of the other (Bubeck, 1995, p. 221). Collins elaborates that responsiveness is a function of perceived need, the intention for

responsive action, and an interpretation of the extent to which the action will indeed fulfil the need of an other (Collins, 2015).

Ethics of care scholars have argued that researchers can also be responsive, also when their data subjects are spread over the world (Barnes & Brannelly, 2015; Held, 2005). For example, Tula Brannelly asks researchers to take the time to gather responses from potential research subjects about their willingness to be subjected to experimentation, and to understand their needs. She notes that trust is likely breached when participants indicate an uncomfortableness with the proposed data collection and changes to their environments (Brannelly, 2018). This is not too different from an informed consent requirement when conducted in person, though it could also be initiated a step before the data collection mechanisms are even designed.

For this thesis, I understand the research theme "responsiveness" as:

1. The responsibility to identify the needs of vulnerable users who are implicated in a context and thus have a relationship with the researcher,
2. The effort to understand these users' needs and interests, and
3. The interpretation and accommodation of their needs in the (technical) research design.

Research theme 4: Emotions and attitudes as moral tools

The previous research theme describes responsiveness as an interrelation of action and thought. This research theme focuses on the thoughts that inform a caring action, which is characterised by both attitudes and emotions. Further, the outcome of a responsive action is judged by the characterisation of the resulting relationship.

The focus on emotions and attitudes in relationships contrasts the ethics of care from other moral theories, which consider emotions and attitudes as interfering with abstract and universal approaches.

Attitudes are the characterisation of an agent's response to a perceived need. Collins describes positive attitudes as promoting, respecting, or revering the interests of another with whom they are in a form of a caring relationship. Further, she states that "some of the most morally important actions and attitudes aim to value, preserve, or promote personal relationships" (Collins, 2015, p. 37). Bubeck describes the application of the ethics of care as requiring "sensitivity to the features of the particular situation as well as imaginative sympathy." She stresses the importance of an attitude that makes room for intuition in guiding decision-making (Bubeck, 1995, p. 221).

In turn, attitudes are informed by emotions, both positive and negative. Ethics of care scholars have put forward a wide range of emotions that guide a moral agent. Collins describes a positive response to another's need to be based on sympathy for the other, a desire to promote the others' interests, including their affects, desires, and decisions (Collins, 2015). Engster and Hamington identify empathy as the key guiding principle (Engster & Hamington, 2015). Negative emotions that have been explored in the ethics of care literature include domination, exploitation, mistrust, hostility, oppression, injustice, and paternalism (Held, 2005; Collins, 2015).

For this thesis, I understand the research theme 'emotions as moral tools' to be descriptive of researchers' attitudes and emotions in the decision-making of their technical design. An appropriate decision reflects the empathy and sympathy of the

researchers towards their unsuspecting data subjects and their sociopolitical contexts. Inappropriate decisions either maintain or further entrench the dominating role of the researcher over the data subject and exploit their unawareness, possibly leading to mistrust of the researcher, research on the Internet, and the research profession in general.

Relevant applications

Literature on practical applications of the ethics of care is somewhat scarce compared to other normative theories. Most scholars have focused on contexts such as healthcare and nursing (e.g. (Molterer, Hoyer, & Steyaert, 2019)), the environment (e.g. (Deacon & Williams, 2004)), and social policy (e.g. (Deacon & Williams, 2004)). Only a handful of papers address law or information and technology law and policy specifically. Three papers are directly relevant for the conceptual framework of this thesis.

(1) In 'Power, care, and the reach of the law', Held argues that the ethics of care offers an alternative or complementary approach to the law for the interpretation of social and human conflict (Held, 2005). The theory takes into account positions of power and privilege and their resulting responsibilities in its consideration of rights and equality. Held builds on Sevenhuijsen, arguing that an ethics of care interpretation ensures that the interests of the more vulnerable in society are not overshadowed in the otherwise abstract and universal reasoning that typically underpins the law.

When judging the relationship between the researcher and their data subjects, power may relate to the superior position of knowledge of the academic and the privilege

of academic freedom that has been granted to them. The vulnerable may be the unsuspecting Internet user who is implicated in the research. Thus, the rights awarded to the abstract autonomous and self-interested person envisioned by the law may not be appropriate to them.

(2) In 'Care, normativity, and the law', Rita Manning states that "care ethics can provide a valuable conceptual and normative resources of many issues in law [...]." However, she ends this sentence stating that "[...] much work needs to be done before care ethics can explicitly play such a role" (Manning, 2015, p. 127). She assesses several judgements in law where the ethics of care may have played a role in legal reasoning.

The most relevant case is the California Supreme Court judgement in *Moore v. Regents of the University of California* in 1990. In short, Mr Moore was diagnosed with a form of leukaemia and sought treatment at the University of California Los Angeles' (UCLA) Medical Centre. The doctor recognised that Mr Moore's tissue samples had commercial and academic value. He requested that Mr Moore makes several trips from his home in Seattle to UCLA, which were not necessary for the medical treatment, though these trips allowed the doctor to collect the necessary tissue. The doctor and the university patented the cell line developed with Mr Moore's tissue and arranged for its commercial exploitation. The Court held that Moore only had a claim against the doctor and UCLA for an informed consent violation, but no property case with regard to the removed bodily tissue.

Manning takes an ethics of care approach in her assessment of the moral nature of the actions in this case. She argues that Mr Moore was vulnerable as a result of his

medical condition and the trust he had put into the relationship with his doctor. The doctor was in a position of care in his relationship with Mr Moore and had a professional duty to make his health the central focus of this relationship.

Manning—in part quoting another critique of this case by Forell and Sortun—describes the actions of the doctor as an abuse of the doctor-patient relationship, a failure to express empathy and attentiveness, and a betrayal involving exploitation and deceit (Forell & Sortun, 2008; Manning, 2015).

(3) In 'Privacy, surveillance, and care ethics', Marilyn Friedman provides an analysis of the tension between the privacy of citizens and the claims of national security benefits resulting from the implementation of massive electronic data collection programs (Friedman, 2015). Friedman argues that national security is a matter of care managed by a government. Citizens cannot individually protect themselves from the type of threats or harm that results from national security lapses.

Interestingly, Friedman asks whether the action (implementing mass electronic surveillance) is a legitimate means of promoting the value of national security. The question regarding the legitimacy of a technological decision to promote a value is relevant for this thesis.

Similarly, though, she concludes that privacy is an important value for individuals, and as such, an important value for the relational aspects of individuals. She initially assigns more weight to the value of privacy. However, given some technological governance additions (e.g. the requirement of probable cause, a judge's warrant, and public accountability after targeted surveillance), she reconciles the two. Her proposed solutions for the implementation of electronic surveillance technology are not novel as they have been explored by many scholars, policymakers, and lawyers.

The ethics of care therefore approach merely strengthens existing policy positions. Further, the paper contains many references to the feminist roots of the ethics of care, including stereotypes of gender roles (e.g. security portrayed as culturally masculine and or "feminized care work" for which the "[p]rotectors may not have the time or energy"). It is unclear what these stereotypes add to the assessment through the lens of the ethics of care, and I do not include them in my analysis throughout this thesis.

(3) Conceptual Development of Internet Research Ethics

Introduction

As Internet access has increased significantly over the last few decades, so has the amount of research conducted on the Internet. This type of research can be categorised into two distinct areas: social science using the Internet as a medium for research and technical Internet research. Researchers and ethics review boards related to both types of Internet research have struggled to apply the Common Rule to their research ethics review procedures.

The previous chapter outlined the regulatory frameworks of research ethics, its concepts, and their legal and philosophical underpinnings. This chapter explores the new challenges posed by Internet research against the concepts of the Common Rule. Thereupon, the first iteration of the conceptual framework is established, which provides the analytical foundation on which the subsequent case study chapters are built.

The first conceptual framework is informed by three parts of this chapter, and summarized in the fourth: The first part introduces (1) Internet research ethics as a

field of academic inquiry, (2) the Internet as a novel medium for research, and (3) the challenges to the governance of digital information generally. The aim of this section is to describe why the application of existing research ethics frameworks may be problematic in the novel research environments enabled by the Internet, where information behaves differently than it did in previous technological eras.

The second section explores how social science researchers have interpreted and debated research ethics concepts in their work related to the Internet. As a field of enquiry, Internet research ethics has developed quite significantly as a review methodology in the social sciences over the last two decades. This provides a fertile ground for this thesis to build on. This section explores how social science scholars in particular have concluded that the binary provisions in research ethics need to be considered on a spectrum or continuum to make them relevant to the dynamic informational environment presented by the Internet. Further, this section explores the extent to which social scientists have been—wittingly or not—applying reasoning based on the ethics of care to their deliberations.

The third section shifts focus to research ethics debates around technical Internet research. While some debate about the ethical challenges in computer science and network engineering exists, this field is much less developed than other social sciences. This section illustrates how technical Internet researchers have struggled to apply the concepts of the Common Rule to their work. There appear to be two reasons for this. First, the disciplines that fall under the umbrella of technical Internet research are not historically accustomed to incorporating considerations regarding humans as data subjects into review procedures. Second, there may be a lack of institutional support requiring or applying research ethics review to technical

Internet research studies. This section repeats the structure of the second section out of necessity. It compares how the disciplines have approached the various concepts of research ethics when reviewing studies that use the Internet as a medium or means for conducting research. Finally, the end of the third section discusses the conceptual gap that this thesis aims to address, namely the appropriateness of applying and interpreting established principles of research ethics in the field of technical Internet research with unsuspecting Internet users as data subjects.

The fourth and final part of this chapter introduces the first iteration of the conceptual framework. The conceptual framework is presented in a table, which provides clarity on how the concepts of Internet research ethics intersect with the research themes of this thesis. The table is reassessed and developed further in the case study chapters.

Ethics review in a new environment

Internet research ethics focuses, as the name suggests, on the ethical issues that arise in research conducted on the Internet. Researchers from several disciplines have found the Internet to be a fruitful space for the study of social interactions and a system that provides interesting technical research challenges. In 1994, Robert Alun Jones gave an early definition of social science Internet research and noted its potential:

“Research in cyberspace - for example, studies of how people navigate, studies of the content or characteristics of electronic messages, studies of system use and transactions generally promises to be an important part of the social and behavioral sciences into the next century” (Jones, 1994, p. 30)

Proposed definitions of Internet research *ethics* evolved as research practices challenged existing frameworks and new ethical tensions arose. Initially, scholars and IRBs applied existing research ethics frameworks to studies conducted on the internet. This made sense intuitively because most research methods, such as online surveys, participant observations, online interviews and focus groups, and discourse analysis in newsgroups and online communities, were transposed from the social sciences. These approaches were incorporated into another commonly cited definition of Internet research ethics:

"IRE is defined as the analysis of ethical issues and application of research ethics principles as they pertain to research conducted on and in the Internet. Internet-based research, broadly defined, is research which utilizes the Internet to collect information through an online tool, such as an online survey; studies about how people use the Internet, e.g., through collecting data and/or examining activities in or on any online environments; and/or, uses of online datasets, databases, or repositories." (Buchanan, & Ess, 2012)³⁶

Internet research ethics has matured somewhat as a field of enquiry over the last decade. The field now operates at the intersection between applied (or practical) ethics and computer ethics (Buchanan, & Ess, 2008). This field draws inspiration from existing research ethics principles and literature as well as ethics traditions in the life sciences (such as medical ethics and bioethics), the humanities, and social sciences (Kraut, et al., 2004; Salganik, 2017). A scholarly debate has since developed about whether new ethical dilemmas emerge when the Internet is used as a venue or tool for research or whether these dilemmas are similar to the challenges faced in other research realms, despite the significant technological influence (Elgesem, 2002; Ess, 2002; Markham & Buchanan, 2012; Walther, 2002). This debate is reflected in a

³⁶ Online publication only. Available at: <https://plato.stanford.edu/entries/ethics-internet-research/>.

more concise and encompassing definition of Internet research presented by Buchanan:

“IRE is defined as the analysis of ethical issues and application of research ethics principles as they pertain to research conducted on and in the Internet.”
(Buchanan, 2011, p. 90)

This is the definition upon which this thesis builds. It captures the application of research ethics principles and frameworks as well as the analysis of emerging ethical issues when research is conducted on or via the Internet.

The Internet as a research domain

It is commonly understood that research practices are shaped by disciplinary domains and the environment within which research is conducted. This section illuminates some of the technical and social characteristics of the Internet and its diffusion in society. This discussion analyses how the Internet has challenged existing information laws, including research ethics frameworks.

Scholarly debate exists as to whether the Internet can and should be considered a new environment for research (Dutton, 2013), or whether the Internet merely allows for new research techniques and methodologies, serving only as a technical extension of social domains (Boellstorff, 2012). This thesis does not aim to contribute to or analyse this discourse. Instead, this thesis notes that new Internet-enabled techniques and methodologies for research in the social sciences have challenged existing research ethics frameworks to some extent. These challenges have led to wide-ranging scholarly discourse, a re-interpretation of existing research rules, and the establishment of new research guidance documents.

The origin of the Internet's unique features

The Internet is characterised by open accessibility, interconnectivity, ubiquity, interactivity, and decentralisation (Berman & Weitzner, 1997; Helbing, 2013). These features have given rise to new technical and social phenomena. However, the nature of the Internet was not necessarily pre-determined nor designed according to a plan. Rather, it was through a confluence of militaristic, scientific, political, and social needs that the Internet was shaped into the global information system that people can access today (Abbate, 2000; Naughton, 2016).

Originally, a limited technical network—called the ARPAnet—connected between 4 and 113 computers based at military and scientific centres, predominantly in the US. While humans did interact on these interconnected networks, the system was predominantly measured and studied as a novel technical phenomenon (see for example (Karn, P. & Partridge, 1987)). The aim of technical research conducted through these networks was typically to yield insights into the performance of networks, understand their structures, model network and application behaviours, and experiment with new types of equipment.

The militaristic and scientific origins of the Internet influenced its initial design values, reflecting the need for survivability and high-capacity performance (Abbate, 2000). One of the most important inventions has been the Transmission Control Protocol/Internet Protocol (TCP/IP), which manages the flow of data on the Internet and, in short, operationalises the end-to-end network principle (Saltzer, Reed, & Clark, 1984). This principle states that the applications and intelligence of the network reside at the edges rather than in the technologies in the middle of the

network. In theory, this makes the Internet an inherently decentralised network (Agre, 2003; Flanagan, Flanagan, C., & Flanagan, 2010). The TCP/IP enables information networks of all types to be interoperable and therefore to converge on one global, interconnected network.

The World Wide Web was built—figuratively—on top of the Internet to connect information sources seamlessly. It was developed by high-energy physicist, Tim Berners-Lee, to encourage the open exchange of scientific data in the late 1980s (Berners-Lee, T. & Fischetti, 2001). The World Wide Web relied mostly on HyperText Markup Language (HTML) to link sources of information and data to each other. The open and interactive nature of the World Wide Web was useful for scientists and enabled a variety of social interactions. When the ARPAnet was decommissioned in the early 1990s, Berners-Lee released his first web browser. This new application allowed for more accessible public and commercial use of the Internet and the World Wide Web (Hart, Reed, R. R., & Bar, 1992; Mowery, D. C. & Simcoe, 2002).

The Internet as a socio-technical system

As the term suggests, a socio-technical system is composed of a technological and a social subsystem, which interact to form a new, hybrid system. A socio-technical system has come to be defined as "complex heterogeneous linkages" among "knowledge, ritual, artifacts, techniques, and activity" (Pfaffenberger, 1992, p, 509). With regard to the Internet, the concept was further developed to include an understanding that "human sociality may extend across both cyber and physical space" (Whitworth, B. & De Moor, 2003, p. 32).

The unique characteristics of the Internet's infrastructure allows for an ever-expanding technical system and an unprecedented information distribution system in terms of scale, scope, and speed as well as open accessibility, malleability, and generativity (Moor, 1985; Zittrain, 2008; Herkert, 2011). The Internet currently interconnects several billion devices and thereby provides access to many different applications (Kurose, J. F. & Ross, 2013).

Due to the theoretically unlimited number of devices that can connect through the Internet, people today experience pervasive connectivity and subsequently communicate, collaborate, and experience or manage their social interactions via the Internet (Azua, 2009). It is estimated that worldwide 51% of people have access to the Internet. More specifically, in the US, 87.9% of the population has Internet access, compared to 85.7% in the European Union. Countries such as the UK and the Netherlands both have an Internet connectivity rate of 94.8% (Stats, 2018). In these regions, the Internet mediates a considerable and increasing amount of human and societal activities. In many parts of the world, the Internet has become deeply embedded into social, political, and economic life (boyd, 2007; Nissenbaum, 2004), displaying the characteristics of a socio-technical system (Flanagin et al., 2010). These technical and social phenomena have created interesting avenues for researchers to explore.

Challenge to information laws

Beyond human interaction, the Internet has had a significant impact on the characteristics of information and on the way information is (or can be) regulated. The accessible, interconnective, ubiquitous, interactive, and decentralised nature of the Internet has enabled the development of a dynamic informational environment

where digital information can be stored, linked, and accessed indefinitely. In theory, infinite copies of information can be made and transmitted across the world efficiently and at a relatively negligible cost (Kaufmann, Manjili, Hildenbrand, Kossmann, D., & Tonder, 2013; Brown, 2013; boyd, 2007). Data storage costs are also decreasing, which means that data can be collected and stored at a large scale with minimal technical limitations (Eynon, Schroeder, R., & Fry, 2009).

The characteristics of information before computers were interconnected differed greatly. Information was previously mostly stored on stand-alone computer terminals, on printed paper, or on other physical carriers. Its distribution and uses were therefore limited—to some extent—to a physical location, and required the physical distribution of carriers. This informational environment is better described as static, as opposed to the dynamic nature of the information environment enabled by the Internet. It is important to make this distinction for this thesis because many of the laws that govern information were written in a time when the dynamic informational environment could not realistically be foreseen.

Information laws—such as those related to data protection, intellectual property, and, indeed, research ethics—therefore rest upon assumptions of a static informational environment. The new characteristics of information or the socio-technical activities enabled by the Internet have had a profound effect on the application of information laws, leading to new power imbalances, inappropriate legal outcomes, or even the criminalisation of a vast number of persons as a result of new activities.³⁷ These challenges have led to regular legislative reviews of

³⁷ See for example: Alexander, I. (2017) “Criminalising Copyright: A Story of Publishers, Pirates and Pieces of Eight”, *The Cambridge Law Journal*, Vol. 66, No. 3.

information laws in many jurisdictions. Several aspects of research ethics laws or frameworks that regulate the use of information for research are prone to the same flaws as other information laws when applied to activities enabled by the Internet.

Some scholars typically argue that not law, but technology design determines how information flows. Joel R. Reidenberg, for example, has stated that the design of the technical information infrastructure sets the de facto rules for how information can be disseminated and thus used. In his legal theory called *Lex Informatica*, he states that:

“[i]n the context of information flows on networks, the technical solutions begin to illustrate that network technology itself imposes rules for the access to and use of information” (Reidenberg, 1997, p. 568).

Similarly, Lawrence Lessig built on *Lex Informatica*, stating that “code is law”, by which he means that a system’s technical design plays a more defining role in shaping information flows than applicable legislation (Lessig, 2006). Nissenbaum further notes that constraints on information flows correspond to the logic of what the technology allows, rather than the social, ethical, and political expectations of its users (Nissenbaum, 2011).

Based on these analyses of the interaction of laws written for a more static informational environment and the dynamic environment enabled by the Internet, another group of scholars have argued that it is difficult, inappropriate, or even futile to apply existing laws to newly enabled activities. Gary Marchant et. al refer to this situation as the “pacing problem”, warning that regulatory processes are too slow to develop adequate regulatory mechanisms. This pace has fuelled concerns about a growing gap between the rules on which society relies to manage its

emerging technologies and socio-technical reality (Marchant, Allenby, & Herkert, 2011).

James Moor refers to the resulting situation as a "conceptual muddle" between technical reality and the law in the books (Moor, 1985). He calls for conceptual frameworks to be developed to analyse the technologies in practice and formulate appropriate policy for action. However, the timing is important. David Collingridge describes the timing dilemma that confronts policy makers, stating that the impact of technologies may be difficult to assess until it has been diffused sufficiently across society. However, changing the parameters of social control (e.g. through reviewed legislation) of a particular technology becomes more difficult it becomes further entrenched in society (Collingridge, 1980). Given the increased attention to the ethics of Internet research and development, I feel that the window for legislative review of technical Internet research ethics has been opened.

Social Science Internet Research Ethics

An increasing amount of social science research is being conducted via or on the Internet. As a result, a variety of groups and networks have come into existence to support this pursuit, such as the Association of Internet Researchers³⁸ and the Network of Internet and Society Research Centres.³⁹ Scholars from a wide variety of disciplines take part in these communities, from disciplines as varied as information science, ethnography, political science, media and cultural studies, literature, law, geography, and many others. The researchers operating on the Internet employ a vast array of traditional and experimental methodologies, including empirical,

³⁸ See <<https://aoir.org/>>.

³⁹ See <<https://networkofcenters.net/>>.

theoretical, controlled, grounded, quantitative, and qualitative methods (Bassett, E. H. & O’Riordan, 2002).

Social science researchers can now automate their data collection on the Internet (or source it from Internet platforms themselves), by continuously collecting and analysing data from around the world rather than being constrained by the limitations of time and space (Dutton & Piper, 2010). Data can now be collected from previously unavailable populations or about types of human behaviour (Eynon et al., 2009). New technologies and methodologies of conducting social science research have raised new and different ethical issues, such as whether Internet users can and do expect their behaviour to be recorded and analysed for research purposes when they interact with others online. It has been suggested that some of these researchers treat the Internet as a social science laboratory (Eynon et al., 2009).

This section first describes two cases from different eras, where social science research on the Internet has led to widespread controversy and criticism. Both cases have sparked interest in the application of research ethics frameworks to Internet research in their time. The subsequent sections discuss how social scientists and ethicists have interpreted and re-conceptualised the concepts of the Common Rule (and other research ethics frameworks) when challenged by social science research on the Internet. Then, the section engages with some scholars’ outspoken opinions about the need for a new research ethics framework. Finally, the section closes with an overview of the emerging research ethics guidelines for Internet researchers in the social sciences, illustrating the efforts that have been undertaken so far to make research ethics concepts more applicable to Internet research.

Controversies

Social science research on the Internet has been plagued with controversies for several decades. This section describes both an early and a recent controversy. These examples set the scene for the following sections, which analyse the ambiguity of research ethics concepts and their resulting academic deliberations.

(1) Rimm (1994) 'Marketing pornography on the information superhighway'

The challenges to research ethics posed by Internet research became apparent initially in the mid 1990s when Marty Rimm—then an undergraduate student at Carnegie Mellon University—published a study titled 'Marketing pornography on the information superhighway' (Rimm, 1994). The aim of the study was to create a typology of online pornography and its viewers. To this end, Rimm analysed the textual descriptions of digital files containing pornography taken from online forums (more specifically, 35 different bulletin board systems) that were not publicly accessible. He further downloaded other information, such as the list of files stored on personal computers of the forums' users and their personal information, including—to the extent available—their names, telephone numbers, addresses, and driver's licenses (Thomas, 1996; Buchanan, 2011). The study gained much attention in the media and was even featured on the front page of Time magazine.⁴⁰

The university and associated researchers quickly distanced themselves from the study when it became widely criticised through the lens of research ethics.⁴¹ Jim Thomas, for example, found the methodology to be excessively invasive into unsuspecting Internet users' lives and labelled Rimm a "peeping cyber-Tom"

⁴⁰ See <<http://content.time.com/time/covers/0,16641,19950703,00.html>>.

⁴¹ The author claims to have consulted "privacy experts" but there is no indication the study was reviewed on the basis of a research ethics framework, such as an institutional review board (IRB).

(Thomas, 1995). Thomas further accused Rimm of misjudging the nature of the collected data as *publicly available*. While the data may have been technically accessible for a person with the necessary technical capabilities, Thomas considered Rimm's activities to be more akin to gaining illicit access without informed consent and to conducting deceptive research without taking traditional precautions (Thomas, 1996).

Buchanan notes that this case "showed a growing uncertainty around human subjects models and Internet research", which forces a reconsideration of fundamental issues of "consent, access, privacy, anonymity and identification, researcher integrity, misuse of data by researchers, representation, limits of 'public' and 'private' data, and ownership of data" as well as the changing parameters of human subjects in Internet research (Buchanan, 2011, p. 89).

Jim Thomas further suggests that a special duty exists for a researcher with the technical capabilities to collect private and (potentially) sensitive data on the Internet. Rather than being allowed to publish this data, Thomas argues that

"[...] it confers upon researchers an obligation to exercise special caution when information is obtained from informants who do not know they are the subjects of a study and are enticed to provide information about third parties who are unaware that information about them is being gathered, studied, and eventually made public." (p. 193)

This statement is central to the aim of this thesis. Given the ability of Internet researchers to collect, process, and disseminate data about research subjects without their knowledge, what is the role of research ethics frameworks to impose an exercise of special caution on researchers? How can this be achieved? This brief case

study demonstrates that Internet researchers are able to create a relationship with unsuspecting Internet users by accessing their computer systems, thus violating their claims to privacy and autonomy by copying and making available the content for other purposes. Further, a negative effect on the relationship with colleagues is observed when the associated researchers distance themselves from the project. The ethics of care methodology can be fruitful in assessing cases in Internet research. The next case illustrates how these practices persist two decades later, with different effects on the lives of unsuspecting research subjects.

(2) Kramer et. al (2014) 'Experimental evidence of massive-scale emotional contagion through social networks'

In January 2012, a Facebook employee and researchers from Cornell University conducted an experiment to discern whether emotions could spread online by studying whether social media users were susceptible to "emotional contagion" (Kramer, Guillory, J. E., & Hancock, 2014). The researchers purposefully presented posts identified as containing either particularly positive or negative emotional language to nearly 700,000 specific users. They then collected data about the posts shared by these targeted users in order to assess whether these contained similarly negative or positive language (depending on what the specific user was shown). This experiment also allowed the researchers to infer whether it was indeed possible to spread emotions via social media manipulation. The researchers concluded that the answer is "yes" and published their findings in the Proceedings of the National Academy of Sciences in 2014.

The researchers at Facebook and Cornell did not seek research ethics review for their study. It is undeniable that the research involved experiments on human subjects,

but researchers and representatives from their organisations argued that they were not required to seek review for a variety of reasons. First, in a letter to two inquiring professors, Facebook claimed that the intent of the study was to improve its products and services and not to contribute to generalisable knowledge.⁴² This argument was easily refuted by their critics who pointed to the published paper in a scientific journal, which makes claims about human behaviour that go beyond Facebook as an online platform (Grimmelmann, 2015). Second, in the same letter, Facebook's representatives argued that the research ethics laws "were not designed to address research conducted under these circumstances." While, strictly speaking, the laws of research ethics were indeed initially developed for the field of biomedical sciences, Grimmelmann points to the common application of these laws to review proposed social science research.

The researchers at Cornell, on the other hand, attempted to circumvent research ethics review based on a self-serving interpretation of the research ethics laws and a technicality. First, Cornell's IRB argued that their researchers were "not directly engaged in human research."⁴³ The IRB was thereby most likely claiming that Facebook was doing the data collection and that its researchers were merely analysing existing data. This reasoning is questionable, especially as the research paper states that all authors were engaged in designing the research (Kramer et al., 2014), which clearly included interaction with participants and a manipulation of their environments. Further, U.S. research ethics laws do not require research ethics review when the study is not federally funded. Since the study was funded by Facebook, the Cornell researchers left the 'federally funded' box unticked and

⁴² See <http://james.grimmelmann.net/files/legal/facebook/FacebookResponse.pdf>.

⁴³ The official statement has been removed from Cornell University's website, but secondary sources cite the original. See for example: <<https://www.theguardian.com/science/head-quarters/2014/jul/01/facebook-cornell-study-emotional-contagion-ethics-breach>>.

proceeded without prior research ethics review. However, a patchwork of U.S. State laws also applies to research that is carried out on American citizens. For example, the applicable laws in California, New York, and Maryland would require this research to be reviewed by an institutional review board, regardless of the funding source (Grimmelmann, 2015). Since Internet research on an unspecified group is likely to include unsuspecting participants from these states, their laws would have therefore required Cornell to seek research ethics review.

The ethics of care is relevant for this case, too. First, beyond the relationship between its users and the social media platform, a relationship is established between the researchers and the unsuspecting Internet users by manipulating the information that is provided to them. The relationship can be characterised as having a negative value for the users, because they are manipulated, and their autonomy is violated. The researchers were not responsive to the users' needs, but instead used them as a means to an end.

Challenges to existing research ethics concepts

The following subsections explore how scholars have understood the extent to which Internet research challenges existing concepts of research ethics frameworks. These existing concepts include *human subjects*, *minimal risk*, *autonomy*, *privacy*, the *public/private distinction*, *anonymity* and *identifiability*, and the *reuse of data*. Each section illustrates how concepts of the research ethics framework, as developed for the biomedical sciences, have evolved as they have been interpreted for social science Internet research. Not every concept has been satisfactorily re-interpreted and therefore may remain somewhat unresolved.

Human subjects

In the literature review, the concept of human subjects of the U.S. research ethics framework was described as a relatively uncomplicated concept. A research participant is either categorised as a human subject (i.e., a living individual about whom identifiable private information is collected through an intervention or an interaction) or they are not considered a human subject. This binary concept has been challenged increasingly in the context of Internet research. The technical reality of online data collection and experimentation has led scholars to suggest an interpretation of the definition provided in the Common Rule that is based on a continuum, rather than a binary.

The main argument to revisit the interpretation of §46.102(f) of the Common Rule is the technical reality and its consequences on the relationship between Internet researchers and their data subjects. Computers and the Internet have disintermediated the traditional relationship between researcher and data subject, creating a geographical and metaphorical distance between the two. Using the Internet, researchers can collect data from widely dispersed populations that may live anywhere in the world at a relatively low cost and in much less time than would previously have been possible given the limitations of the physical world (Frankel, M. S. & Siang, 1999; Keller, H. E. & Lee, 2003). The data subject becomes a 'distant other' in terms of the ethics of care. Both the Rimm and Facebook cases illustrated that identifiable private information can be collected about Internet users through an intervention or (technical) interaction without the data subject's explicit permission or even knowledge. Human subjects can also be implicated in Internet research in a variety of other ways, which are more subtle than would have previously been possible offline (Wright, 2006). For example, the Internet offers the technical ability

to collect and combine data that seem to be available freely and publicly, such as peoples' online conversations in forums or their browsing habits (Frankel, M. S. & Siang, 1999).

This experiential distance to others, increases when the persons producing the data collected for research are not easily identifiable, such as in online chat rooms or—on a more technical level—in the analysis of data based on IP addresses. The perception of the data subject as a human subject is relevant for Internet research ethics because the classification of a human subject in a research study is what typically triggers a research ethics review. However, scholars have noted that as the distance increases, data subjects are less likely to be considered human subjects by researchers or ethics committees (Aycock, Buchanan, Dexter, & Dittrich, 2011; Buchanan, Aycock, Dexter, Dittrich, D., & Hvizdak, 2011). Eynon et. al note that the distance between a researcher and data subject makes it more complicated for a researcher to assess the risk of harm on individual research subjects, as their social environment remains unknown. Furthermore, this distance complicates the ability of the researcher to judge an emerging situation and intervene to halt data collection and experimentation when something goes wrong (Eynon et al., 2009). This lack of information about the other, complicates any efforts at responsiveness.

Buchanan explains that as the nature of research and the boundaries between its actors are being disrupted by the Internet, it makes sense to ask how some of the concepts of research ethics are changing as well. She concludes that due to its socio-technical nature, most research on or via the Internet implicates humans in some way. Therefore, she suggests that we think of research methodologies involving human subjects as a continuum. On the one end, humans might not face any risk at

all (e.g. aggregated and decontextualised data analysis devoid of identifiable human connection). Further along the continuum, data subjects may not be directly identifiable, but a range of ethical issues may arise, even to the surprise of the researcher. On the other end of the continuum, humans can be directly identified or their intimate or personal information can be revealed (e.g. digital ethnography) (Buchanan, 2011). This conception builds on a suggestion by Bruckman, who has put forth that it may be more useful to think of the identifiability of human subjects in research as a continuum, rather than a binary (Bruckman, 2002).

Minimal risk

As with many other concepts in research ethics, the pivotal requirement of ‘minimal risk’ changes or is challenged when applied to online experiments (Kraut et al., 2004). The Internet enables the potentially unrestrained data collection from not-necessarily identified data subjects, across many geographic locations, and multiple contexts. This data collection differs from offline research, during which physical and personal interactions with data subjects allow the researcher to understand the subtleties of their expectations of the research. For example, in the context of qualitative Internet research, Schrum asks:

“How then does one study these new electronic communities, given that the researcher cannot easily share a meal? How does one find an ethical way into local context when that local context is scattered around the world?” (Schrum, 1995, p. 313).

The minimal risk requirement is described in the previous chapter as referring to harmful psycho-social consequences. The magnitude of these consequences and probability of occurrence are described as less or equal to the experiences of a

normal person in their everyday life. The variety of ways in which humans can be implicated in research, however, makes the calculation of risks of harm and benefits more complex and even creates new risks (Boehlefeld, 1996; Dittrich et al., 2012). Internet research can potentially implicate people from a variety of backgrounds and a wide range of geographic locations, thereby complicating context-oriented deliberations. The social, political, and economic environments in which these Internet users live defines to a large extent the risks they face in their daily lives. For example, risks from engaging with research—wittingly or not—can arise as a result of specific local informational sensitivities, applicable laws, and power relations to which data subjects may be subject to. Only when the local context and expectations of individual data subjects are known and accounted for, can researchers meaningfully understand the threshold for minimal risk, and therefore be responsive to their data subjects' needs. Furthermore, the risks to unwilling participants may only become apparent when the results have been published, which may be too late to implement any meaningful mitigation strategies (King, 1996).

Reflecting on some of these complications, Cappuro and Pingel, too, advocate for an ethics of care for social science Internet researchers, though without much substantiation (Capurro, R. & Pingel, 2002). Some scholars have encouraged Internet researchers to inadvertently operationalise an ethics of care by understanding the expectations of data subjects (Elgesem, 2002). For example, if data subjects are part of a defined online community, Pittenger suggests a contextually-oriented deliberation. Researchers should seek out any established rules, etiquette, or guidelines that apply to participants and outsiders (Pittenger, 2003). It seems, thus, that Internet researchers should not use a single, universal 'minimal risk' standard. Instead, there appears to be an expectation that researchers ensure their research studies do not

impose greater than minimal risk on research participants in their own environments. Again, it would make sense to think of the minimal risk standard as a flexible concept, or a continuum, which allows for the evaluation of different levels of risk for participants in different environments.

Autonomy

The previous chapter identified the concept of autonomy as central to the aim of research ethics. Researchers can have a certain amount of power over individual data subjects, and this power can be rebalanced by research ethics review procedures that impose obligations such as transparency requirements, the gathering of informed consent, and debriefing after deception research. Safeguarding research participants' autonomy enables them to choose whether to participate in a research study and withdraw for whatever reason. It also protects vulnerable groups, such as participants in deception research. Research conducted on the Internet changes some assumptions and considerations, though, as it enables researchers to collect data about participants from a distance, invisibly, and without their consent. This section assesses how the safeguarding of autonomy has been discussed in social science Internet research ethics literature, specifically with regard to informed consent, vulnerable persons, and deception research.

Informed consent

The previous chapter noted that some scholars use the informed consent procedure to operationalise research participants' autonomy and the requirement of respect for participants. Other scholars challenge the extent to which an informed consent procedure is sufficient to establish respect for individual autonomy. Discussions about informed consent in social science Internet research have revealed that gathering informed consent from data subjects on the Internet raises a range of new

issues, giving weight to the criticism that informed consent may indeed not be sufficient to operationalise autonomy or the respect for persons. In summary, scholars argue that gaining informed consent both in terms of participants understanding what they are consenting to (Varnhagen et al., 2005) and the extent to which individuals are able or competent to give informed consent (Kraut et al., 2004; Pittenger, 2003) may be more challenging in the online environment (Eynon et al., 2009).

Scholars note that one of the main challenges for the legitimacy of informed consent may be that the identities of both researchers and data subjects could be unknown to each other. For example, members of an online community may use pseudonyms that obfuscate their identities, or they may only be separable by unique IP addresses in a database. Similarly, the researcher may hide their identity, even when asking for the informed consent of data subjects (e.g., in the case of deception research). Informed consent may establish a better, more clear relationship between researchers and their data subjects, beyond a mere connection.

When the identity of the human subject, their geographic location, and methods of communication between parties are unclear after consent has been gathered, Frankel and Siang ask “[...] how valid is such consent when the age, competency, or comprehension of the potential subject is unknown?” (Frankel, M. S. & Siang, 1999, p. 10). Researchers would need to be able to notify Internet users of their right to withdraw from the study, but would be unable to do so if the identities of individual data subjects are unknown (Kenneally, 2015a; Peden, B. F. & Flashinski, 2004). The ability to be responsive is diminished when identities are unknown within a relationship

Finally, a challenge exists when collecting research data from a group of Internet users, such as a virtual community. Obtaining informed consent from all users of a group would likely be difficult to operationalise. This is partly because membership of a group under study can be constantly in flux, so not all members may be aware of an ongoing study, and some may object to participating if presented with an option (Keller, H. E. & Lee, 2003). Furthermore, even if some do not consent, their interactions with others who have consented could reveal information about them, too. The consent of some users does not give assurance that all members of the community have consented or have knowledge of the data collection for research purposes (Guba & Lincoln, 1989; King, 1996). Researchers would therefore likely find it impossible to clearly inform all (possibly hugely diverse) participants about the expected risks of harms, the benefits, and the processes of the research in a meaningful way and in a single document that is relevant to all potential users in their specific contexts (Buchanan, & Ess, 2008).

Vulnerable persons

Literature on the inclusion of vulnerable people in Internet research has been scarce. Some scholars have explicitly considered the effects on vulnerable groups, but typically refer to ideas, concepts, and guidelines regarding offline research with vulnerable groups and apply them to their online studies. For example, Siobhan Sharkey et al. conducted a study on a self-harm discussion forum and made the case for heightened research ethics responsibility based on traditional vulnerable groups being at increased risk (Sharkey et al., 2011). They specifically referred to *Ethics-as-Process* models developed in the domain of health research, which enables researchers to be more reactive and responsive to emerging risks of harm for

vulnerable research participants (Cutcliffe, J. R. & Ramcharan, 2002; Ramcharan, P. & Cutcliffe, 2001). Sharkey et al.'s research targeted a specific community, whose participants could be assumed to be more vulnerable than the average random population. In cases where the vulnerability of the population is unknown, Eynon et al. have raised concerns that the problems of accurately identifying distant, other research subjects in Internet research may lead to researchers being unaware that the population contains vulnerable groups that would require more care (Eynon et al., 2009). Since the Internet does not have a so-called 'identity layer', most Internet researchers would face challenges in confirming the identity of their data subjects. Finally, the Association of Internet Researchers guidelines of 2012—which are discussed in a further sub-section—mention the "greater vulnerability of [...] participants, the greater obligation of the researcher to protect [...]" as its first key principle (Markham & Buchanan, 2012, p. 4). While the guidelines raise the issue of vulnerability in several sections, the focus is mostly on minors participating without parental consent.

Deception

Deception research can be conducted relatively easily on the Internet. Researchers may create a technical relationship by exploiting the invisibility of computer code to collect data, or by collecting data about behaviours through other unobtrusive or covert methods. Alternatively, researchers can pose as a member of a community by creating a misleading online identity (Frankel, M. S. & Siang, 1999). It is important to note, however, that the covert collection of data does not necessarily involve deception, but rather typically a lack of informed consent (King, 1996).

One of the earliest mentions of deception research in Internet research was in Elizabeth Reid's reflection on her own qualitative study into online communities. Reid decided it was practically impossible to disclose her presence as a researcher to all participants in the community under investigation (Reid, 1996). She did add some labels to the description of her avatar (ie. her virtual persona in graphical format), such as a notebook and a tape recorder, to give some indication that she was logging the conversations. She mentions herself, though, that the efficacy of this approach to transparency is questionable. Therefore, she also asked for permission from the administrators of the particular community (assuming they would be in a position to decide for their users as a proxy), and sought permission to publish quotes from participants.

After publishing her research, the online community Reid had studied attracted some attention from aspiring members and other social scientists who wanted to study the community as well. The administrators had to tighten the security of the group, which further changed its internal dynamics. While Reid states that she does not think that her research was the sole contributor to the changes in the community, it did play a part. One of the administrators described the effect on the environment of the community as forcing "the members of the system to trade a 'freedom to' for a 'freedom from' [...]," (p. 174) indicating that proactive measures now needed to be taken to allow participants to communicate freely. The result of her deception research prompted Reid to reflect on her assumptions. She stated:

"I do not think that researchers should necessarily always carry the onus of responsibility to second-guess their subjects' reasons for consent. Such an expectation of paternalistic empathy would be unreasonable." (p. 172)

and:

"[...] I now believe that I should have insisted on anonymising all quotations, even if this meant that in some cases I would be refused permission to quote." (Reid, 1996, p. 173)

Required debriefing and retroactive informed consent from data subjects, which are a form of responsiveness, pose a challenge to the legitimacy of deception research in Internet research. For example, when deceiving entire communities or groups of Internet users, it is difficult or infeasible to communicate a debrief effectively to all deceived participants. The identification of each participant would require that each individual's personal information be gathered before they leaves the virtual space within which the research is conducted (Keller, H. E. & Lee, 2003). Pittenger argues that a meaningful debrief would "would require some extraordinary logistical oversight to ensure the protections for the participants" (Pittenger, 2003, p. 56).

Privacy

The Internet allows researchers to collect unprecedented amounts of social research data through technical means. The vast amount of data being produced have led some social scientists to note the potential of the Internet to end social scientific theorising (Anderson, 2008; Kitchin, 2014; Mazzocchi, 2015). One researcher – Nicholas Christakis – was quoted as saying:

"We're on the cusp of a new way of doing social science... Our predecessors could only dream of the kind of data we now have" (Rosenbloom, 2007)⁴⁴

The sources of research data can vary, from scraping Internet platforms and forums to experimentation with peoples' devices such as their computers or mobile phones. Scholars note, however, that these online identities and devices are becoming

⁴⁴ Online publication only. Available at: <https://www.nytimes.com/2007/12/17/style/17facebook.html>.

extensions of natural persons and their environments (Clark, A. & Chalmers, 1998). Online platforms invite users to share a vast amount of data about themselves, and peoples' devices are typically used to store personal notes, pictures, videos, and communications. An internet user's identity is intertwined with their devices due to the increasing sociotechnical convergence of the human experience through computer networks. Therefore, any outside interaction with the data contained in or transmitted through these platforms and devices may be seen as an invasion of the owners' privacy (Buchanan, 2011; Roberts, 2014).

Critics state that people can be covertly recorded in the physical world and information can be gleaned about them from various sources. Eynon et al. note, however, that the capabilities for recording, reproducing, and analysing online interactions on the Internet are far more powerful than in the physical world (Eynon et al., 2009). In fact, by merely using the Internet, users typically leave behind a data trail that can be collected by a large range of persons at a later date. Michael Zimmer notes that:

“[...] merely having one's personal information stripped from the intended sphere of the social networking profile, and amassed into a database for external review becomes an affront to the subjects' human dignity and their ability to control the flow of their personal information.” (Zimmer, 2010, p. 321)

This concern was further developed in Helen Nissenbaum's theory of contextual integrity, which considers the appropriateness of new data flows through the lens of the social context within which information was revealed (Nissenbaum, 2004; Nissenbaum, 2009). Relatedly, Eynon et al. connect the ability to collect data globally to the issue of due consideration of the risk of harm for data subjects (Eynon et al., 2009). Data collection is potentially disconnected from the context within which it

was first revealed, by which the resulting dataset loses all the contextual sensitivities which a researcher may need to take into account.

While the Internet does, indeed, enable a new way of conducting social science, it raises a range of new issues. New and expanded capabilities for data collection typically produce new and expanded information privacy considerations. In social science Internet research, the concept of privacy has been discussed most prominently in three different ways. First, the Internet has blurred the distinction between *publicly available* and *private* data. Second, the dichotomous distinction between individuals being *identifiable* or *anonymous* no longer holds true. Third, the reuse of research data has been questioned, especially with regard to large datasets. These topics are addressed in the next subsections.

Public/private

Existing research ethics frameworks typically draw a distinction between data collected from public and private spaces. Data collected in public spaces is perceived to be a lesser infringement of privacy than data collected from private spaces (Elgesem, 2002; Sveningsson, 2001). The Rimm case illustrates that while a vast amount of information may be technically available and accessible via the Internet, it may not have been intended to be publicly available for research purposes or otherwise. A difference exists between publicly *accessible* information, which is protected by various data protection and computer security laws, and publicly distributed information, which is *intended for public use* (Frankel, M. S. & Siang, 1999).

Christina Allen noted that conceptions of public and private can be blurred because both types of spaces can exist within the same larger venue on the Internet (Allen,

1996). As a result, researchers need to seek analogies to justify why their online data can be compared to public offline spaces, such as 'public chatrooms' being equated to 'public squares'. However, analogies are not necessarily robust. Collecting data from an online chatroom could also be compared with tapping telephone lines (Keller, H. E. & Lee, 2003), which would be considered inappropriate for research and also simply unlawful. Ultimately, the Internet is neither a public square nor a telephone network, but a new medium, for which new rules and norms need to be developed (Bruckman, 2002). It is considered to be more useful to understand the perceptions and expectations of privacy by potential data subjects of the space where data will be collected for the research, rather than relying on offline analogies (Pittenger, 2003).

In another example, Storm A. King raised the issue of unsuspecting Internet users being monitored for research purposes without ever having consented to the collection of their data. In his 1996 paper reflecting on the surveillance-like practices of conducting qualitative psychological research in virtual self-help communities, King published a quote from a member of an online self-help community, through which members found out they were being studied:

"When I joined this, I thought it would be a *support* group, not a fishbowl for a bunch of guinea pigs. I certainly don't feel at this point that it is a 'safe environment' as a support group is supposed to be, and I will not open myself up to be dissected by students or scientists." (King, 1996, p. 122)

King notes that it is a false assumption to hold that anyone publishing a note in publicly accessible forums would know that their writings could be collected by scientists who are interested in studying the underlying psychology. A feeling of being violated may result when information—that an unsuspecting Internet user did

not consider to be publicly available—is published (King, 1996). King recommends that researchers adopt an attitude to assess the need for ethical review based on the concepts of group accessibility and perceived privacy (King, 1996). The two concepts are linked and mutually reinforcing. The less accessible a group or communication platform is, the more members will perceive themselves as acting in a private space. This approach demonstrates a responsive action by a researcher with an attitude to preserve and further promote the needs of their data subjects with the research relationship.

Other scholars agree, urging researchers to consider the perceptions and subsequent behaviours of people based on what they consider to be private communication or information (Elgesem, 2002). Buchanan and Ess even argue that “[t]he expectations of privacy held by those members will dictate the role a researcher may occupy there” (Buchanan, & Ess, 2008p. 280). This implies a move away from the dichotomous decision-tree type distinction between public and private information, towards a spectrum of user expectations.

Thomas, the most prolific critic of the Rimm study, further argued that the ability to acquire private and potentially sensitive data does not give researchers the right to collect or publish this data. Rather, Thomas argues for the special responsibility of researchers who make use of unsuspecting data subjects. He has argued that this knowledge and technical ability

“[...] confers upon researchers an obligation to exercise special caution when information is obtained from informants who do not know they are the subjects of a study and are enticed to provide information about third parties who are

unaware that information about them is being gathered, studied, and eventually made public” (Thomas, 1995, p. 9)

Analogies to offline, physical spaces, where peoples' behaviour can be observed and research data can be recorded, can often be challenged as false comparisons (Zimmer, 2010). The Internet allows vast and simultaneous observation of people in diverse geographic locations, while offering only a more limited ability to discern the identity and other important—only physically observable—variables that may contribute to an understanding of an individual's behaviour.

Anonymity and identifiability

Research ethics frameworks typically consider the use of exclusively anonymised research data to not be a potential risk of harm to data subjects. In practice, however, full anonymisation or de-identification is difficult to achieve. Social science researchers have found that expressions by Internet users published in qualitative research can be retraced to their authors through simple Internet searches (Fiesler, C. & Proferes, 2018). Furthermore, Eynon et al. note that the perceived anonymity Internet users have may lead to them disclosing more information than they would have done in a physical face-to-face interaction (Eynon et al., 2009). The sub-section on ‘Anonymity and Identifiability’ in the section on ‘Technical Internet Research Ethics’ of this chapter goes into more detail because anonymity has been discussed to a greater extent in computer science. As with many other concepts of research ethics, Internet researchers found that it is more effective to consider anonymity and identifiability on a continuum rather than a binary (Bruckman, 2002). By being responsive to the needs that are fulfilled by maintaining the anonymity of data subjects, researchers can become more understanding of the human lives that are hidden behind or within their datasets.

Reuse

Collected datasets for research can be shared, distributed, or otherwise accessed by third-party researchers. Not only existing datasets compiled for research, but also many existing digital datasets can be used for research through digital technologies. The utility of existing data has vastly increased with the use of the Internet. Michael Zimmer illustrates the problem of reusing data for research with a study that collected data from a group of Harvard University students' Facebook profiles and combined the information with data collected by the university's admissions office (See (Lewis, Kaufman, Gonzalez, Wimmer, A., & Christakis, 2008)). Zimmer explains that students had provided personal information to the university's admissions office for various administrative purposes and shared their data with Facebook primarily to connect with their friends. Neither transaction had the purpose of aiding researchers, and neither was a relationship to scholarship intended. The researchers did, however, collect, aggregate, and analyse this data and then distributed it to third parties without the consent of data subjects. Zimmer argues that this use of data by researchers was unauthorised and was a violation of the data subjects' privacy (Zimmer, 2010). While the utility of the data—scientific or otherwise—can be increased if it can be reused for research purposes, laws and ethical requirements may stand in the way of its reuse. Data protection and information privacy laws may restrict the use of data for research entirely, while norms such as 'contextual integrity' discourage the use of data disclosed for one purpose in another context (Nissenbaum, 2004).

Existing social science Internet research ethics guidelines

Beyond formal laws, such as the U.S. Common Rule, a number of research ethics guidelines, specifically for Internet research, have been developed by scholarly associations, organisations, and institutions. These are attempts to fill the policy vacuum left by challenges to existing research ethics frameworks posed by social science research on the Internet. One set of guidelines has been developed specifically for social scientists who conduct research on the Internet, while other, more domain-specific guidelines have been updated to accommodate the increasing amount of Internet research in particular disciplines. The content of these guidelines are mostly based on the themes raised and literature referenced in this conceptual framework chapter. This section comments briefly on the approaches and themes covered in these guidelines.

The most prominent guide is the AoIR's 'Ethical Decision-Making and Internet Research'. The first version was published in 2002, and then, it was updated in 2012 (Ess, 2002; Markham & Buchanan, 2012). The guidelines are being revised at the time of writing this thesis as well. They were initially developed "from the bottom up—i.e., out of the day-to-day practices of researchers in a wide range of disciplines, countries and contexts" (p. 2) to accommodate for the emerging issues faced by researchers who conduct research on the Internet. They focus on several topics, which have emerged in this conceptual analysis. For example, some of the central themes are understanding that the concept of harm may vary with the specific context of the Internet user; considering the vulnerability of data subjects and groups on a continuum with increasing responsibilities for the researcher; clarifying that data collected on the Internet is related to a human; and positioning ethics review as an iterative and deliberative process as research progresses through phases, rather

than a one-time check. The guidelines are mostly informed by experiences in the social sciences, such as the employment of ethnographic techniques in online forums (Thelwall & Stuart, 2006; Eynon, Fry, & Schroeder, 2008; Dutton & Piper, 2010).

Various academic disciplines and institutions maintain their own specific research ethics policies, some of which have become well-established in offline research. As researchers in these disciplines began to conduct some (or all) of their research on the Internet, the need to update research ethics policies became apparent. The new provisions for Internet research ethics are typically included as updated to existing policies or as part of an additional document created to complement an existing policy. Either way, the new provisions for Internet research ethics are fitted into the existing structure, which is typically based on the human subject model of research ethics. The next few paragraphs analyse the focus of some of these documents.

In the US, the prominent American Psychological Association (APA) produced a report in 2002, which provides a set of suggestions for researchers and IRBs when confronted with research on the Internet. The report focuses on the identifiability and anonymity of data subjects, whether Internet communications are to be considered public and therefore can be observed by researchers, and the extent to which information published online can be considered an existing public record that researchers are free to use. The report states that the interpretation of existing research ethics concepts are not clear cut, and puts forth the hope that “[...] in the future, better education about the scope and access of information available on the Internet will align the assumptions of various parties more closely with each other” (APA, 2002)⁴⁵.

⁴⁵ Online publication. Available at: <https://www.apa.org/science/leadership/bsa/internet/internet-report>.

The British Psychological Society (BPS) published a document in 2017 titled 'Ethics Guidelines for Internet-mediated Research', which supplements its core ethics policy called the 'Code of Human Research Ethics'. The authors state that the Internet extends opportunities for research, creating "additional, and sometimes non-obvious, complexities around adherence to ethics principles" (Society, 2017, p. 1). The report re-interprets the BPS' initial principles (respect for the autonomy, privacy and dignity of individuals and communities, scientific integrity, social responsibility, and maximisation of benefit and minimisation of harm) in the context of Internet research. The focus is on concepts discussed in this chapter, such as the public/private distinction, confidentiality, consent, withdrawal, and debriefing after deception. The document further discusses some strategies that can be used to mitigate harms and maximise the benefits of psychological research on the Internet.

Several other disciplinary-specific Internet research guidelines or updates exist in disciplines, ranging from educational research⁴⁶ to criminology⁴⁷ and beyond. Most of these are similar in content, scope, and detail to the psychology-specific guidelines. It is relevant to mention two more examples for the purpose of this thesis: the Economic and Social Research Council's (ESRC) research ethics principles section on Internet-mediated research and the University of Oxford's Central University Research Ethics Committee (CUREC) Best Practice Guidance document on Internet-based Research.

⁴⁶ See footnote 7 in <https://www.bera.ac.uk/wp-content/uploads/2018/06/BERA-Ethical-Guidelines-for-Educational-Research_4thEdn_2018.pdf>.

⁴⁷ See point (4) v in <<http://www.britisoccrim.org/docs/CodeofEthics.pdf>>.

The ESRC is one of the seven research councils in the UK that coordinates and funds research in a number of disciplines. Its website presents a wealth of information about research ethics requirements and expectations, including the functions of research ethics committees. The section on Internet-mediated research briefly addresses issues around privacy, informed consent, identity, and deception research. However, the section is short, it encourages the reader to use AoIR guidelines.⁴⁸

The University of Oxford is a good example of an institution that has made efforts to accommodate research ethics issues that arise in practice. Its research ethics and CUREC website guides researchers towards the appropriate supporting documents in its Best Practice Guidance section.⁴⁹ The Internet-based Research Best Practice Guidance document focuses on methodologies from the social sciences, such as surveys or online text analysis, among a few others. Interestingly, the guidelines state that while informed consent is a key issue, gaining it is “often impractical, and the research must instead be ethically justified via public interest arguments” (p.1). While the document does not elaborate further on what a viable public interest argument may be, it does indicate that the underlying ethical motive is a utilitarian rather than rights-based, virtue, or care-based approach to research ethics.

Technical Internet research ethics

The previous section presented an overview of how the community of social scientists conducting experiments or collecting research data on the Internet have reflected on the ethical nature of their work. The technical Internet research

⁴⁸ See the section here: <<https://esrc.ukri.org/funding/guidance-for-applicants/research-ethics/frequently-raised-topics/Internet-mediated-research/>>.

⁴⁹ See <<https://researchsupport.admin.ox.ac.uk/governance/ethics>>.

community has reflected on the research ethics of their work as well, though the body of work has been applied less, is less clear, and is less well studied compared to the work of social scientists. While the technical Internet research community can learn from the experiences of the social scientists, the disciplines differ enough to complicate a direct transplantation of thought across disciplines.

The main difference between the disciplines is the objects of study. For example, Denning et al. proposed a definition of computer science as “the systematic study of algorithmic processes—their theory, analysis, design, efficiency, implementation, and application—that describe and transform information” (Denning, et al., 1989, p. 12). Relevant sub-disciplines of computer science for the focus on technical Internet research of this thesis are the related fields of computer security and network measurement. Computer security is concerned with the protection of information technology systems and involves the study of subjects such as malicious software, denial of service attacks, and the exploitation of software vulnerabilities (Aycock, Buchanan, Dexter, & Dittrich, 2011). The Internet measurement discipline studies—in short—how the various decentralised components and entities that comprise the Internet behave and interact with one another (Crovella, M. & Krishnamurthy, 2006; Paxson, Mahdavi, Adams, A., & Mathis, 1998). Crandall et al. note that the objects of analysis for these disciplines are not the humans using the Internet, but rather the computers and networks that allow the Internet to operate (Crandall, Crete-Nishihata, M., & Knockel, 2015). Research in these two sub-disciplines is generally assumed to be beneficial for the functioning of the Internet and has begun to involve more human subjects as data points, wittingly or not.

However, several scholars have noted that the practices and language of technical Internet research do not fit into the existing research ethics structures that were developed for the biomedical and behavioural sciences (Aycock, Buchanan, Dexter, & Dittrich, 2011; Buchanan et al., 2011). Furthermore, Metcalf and Crawford suggest that the academic communities that conduct technical Internet research have not traditionally considered themselves to be engaged with human subject research in the first place due to the technical focus of their work (Metcalf, J. & Crawford, 2016). Indeed, Partridge and Allmann reflect on the use of language in their disciplines, which predates the widespread use and diffusion of the Internet in society. The concepts used by computer scientists do not consider humans as end users of systems or data points, but focus on the technology (Partridge, C. & Allman, 2016).

Due to the increasing sociotechnical nature of the Internet, however, several scholars have noted the inadequacy of this approach and therefore highlight the importance of ethical reflection on technical Internet research. The technical knowledge required to understand the impact of this type of research may even make these challenges more pertinent. For example, Internet users are typically not knowledgeable of the inner workings of applications, their devices, and the networks they connect to, nor the data flows they initiate (Grinter, Edwards, Newman, M. W., & Ducheneaut, 2005; Kang, Dabbish, Fruchter, N., & Kiesler, 2015; Poole, Chetty, M., Grinter, R. E., & Edwards, 2008). Highly technical researchers or engineers who perform technical Internet research do understand the inner working of devices, networks, and resulting information flows better than most of their data subjects. Relationships are therefore established by technical means. In fact, the technical choices made in the design of technologies or experiments can have direct ethical, societal, and political consequences for users and data subjects, which they may not be aware of

(Spiekermann, 2015). For example, it is possible to control and manipulate information flows on people's devices in ways that are invisible to their owners. Matt Salganik aptly describes this as the "ability to do things to people without their consent or even awareness" (Salganik, 2017, section 6.3). Accordingly, in a paper reflecting on technical research and information filtering, Wright et al. note:

"[...] we see a number of serious legal and ethical concerns that must be thoroughly considered in order to undertake broad-scale research of this nature." (Wright, De Souza, & Brown, 2011, p. 5)

Several institutions, such as the Association for Computing Machinery and the Dutch Academy of Sciences, are raising general awareness that the powers of technical Internet researchers over society and individuals must be kept in check by imposing responsibilities and moral obligations on researchers who design technologies or experiments (Gotterbarn, Bruckman, Flick, Miller, K., & Wolf, 2018; KNAW, 2016). However, several scholars note that most research ethics review boards do not have the technical expertise or training to review proposed research and impose these responsibilities (Buchanan, E. A. & Ess, 2009; Hall, T. & Flynn, 2001). Partridge and Allman note that, indeed, U.S. Institutional Review Boards typically consider Internet measurement research to be exempt from research ethics review, even if data are collected from individuals' devices. Instead, they note that conference programme committees at computer science conferences are typically the first organisations that review papers, which include reflection of the ethics of research. They note, however, that these review boards do not necessarily make use of the existing research ethics frameworks and have rather concluded that "[t]he measurement community needs to find its own answers" (Partridge, C. & Allman, 2016, p. 62). Similar statements have been made in regard to computer security, such

as in Phillip Rogaway's call to seriously consider the social impact of cryptography. He stated:

"I plead for a reinvention of our disciplinary culture to attend not only to puzzles and math, but, also, to the societal implications of our work."

(Rogaway, 2015, p. 1)

This statement is interesting not only because it displays a moral awakening by computer scientists, but also because it appreciates the significance of the technical Internet researchers' work in establishing relationships with wider society. Research ethics literature and frameworks are typically concerned with the relationship between the researcher and the researched. The broader impacts on society are not necessarily a consideration in research ethics frameworks.

This section reviews some of the challenges, proposed solutions, and proposed reinterpretations of concepts in existing research ethics frameworks for technical Internet research. The research questions are presented after this section, once the gaps in knowledge are made clear.

Research ethics guidelines

Several projects and initiatives have attempted to create ethics guidelines for the computer profession and technical Internet research specifically. This section introduces three such documents.

The Association of Computing Machinery (ACM)'s Code of Ethics and Professional Conduct is possibly the most well-known ethics guide for computer scientists or the document with the widest coverage (Gotterbarn et al., 2018). The Code was initially

adopted in 1992⁵⁰ and was recently revised in 2018.⁵¹ The document provides a high-level overview of some professional ethics requirements common in other disciplines, such as the requirement of competence, abidance of applicable laws, and the fostering of public awareness. In addition, the Code provides some specific guidance for computing professionals, including harm avoidance and non-discrimination when developing and integrating systems into society. Some of the Code's requirements are vague. For example, they ask computing professionals to use their skills "for the benefit of society" or to "minimise negative consequences." The most relevant part for this thesis is section 1.6, which requires computing professionals to respect privacy. The Code states that, indeed, computing technology

"[...] enables the collection, monitoring, and exchange of personal information quickly, inexpensively, and often without full knowledge of the people affected."

The Code does recommend that computing professionals obtain the informed consent of data subjects. More specifically, the Code prohibits the reuse of data collected for one purpose without the data subject's consent. Furthermore, the Code requires that precautions be taken to prevent the re-identification of the individuals behind anonymised data or even the collection of unauthorised data.

The Code seems to be designed to address a number of high-level concerns, though it stops short in terms of its address of the specific challenges a technical Internet researcher may face. As a result, some of the concepts or requirements, including those regarding privacy, can be interpreted in a variety of ways. In fact, a recent

⁵⁰ See 1992 version: <<https://ethics.acm.org/code-of-ethics/previous-versions/1992-acm-code/>>.

⁵¹ See the revision process: <<https://ethics.acm.org/code-of-ethics/code-2018/>> And the final guidelines: <<https://ethics.acm.org/>>.

study of the Code concluded that “[d]espite its stated goal, we found no evidence that the ACM code of ethics influences ethical decision making” (McNamara, Smith, J., & Murphy-Hill, 2018, p. 4).

Another set of guidelines, titled ‘The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research’, was developed in 2012. These guidelines respond to the need for ethics guidance by technical Internet researchers. The report interprets the principles of the Belmont Report for technical Internet researchers who work in areas such as Internet measurement and information security. It is interesting to note that the authors did not explicitly interpret existing research ethics frameworks such as the Common Rule. While the Menlo Report is an important intellectual exercise for the assessment of the Belmont Report in light of new technologies, it is not directly useful—at least not in the US—for researchers who are faced with questions based on the Common Rule by their institutional review boards. The Menlo report has been applied in some technical Internet research papers (e.g., (Khan, M. T. & Kanich, 2015)), though some of the authors of the Report have concluded themselves that ‘the gap between such principles and actionable guidance for the ethical conduct of ICT research is large’ (Bailey, Kenneally, E., & Dittrich, 2012).

Finally, as mentioned in the introduction chapter, I co-authored a set of guidelines titled the ‘Networked Systems Ethics – Guidelines’. The primary aim of the project was to “give substance to the discussion on the ethical, legal and policy issues of networked systems research”. However, possibly more important at the time, the secondary aim was to “build a multidisciplinary network of experts in academia, government, NGOs and business from diverging theoretical and practical

disciplines". Several computer science conferences have referred authors and programme committees to these guidelines to guide their thinking on research ethics. The end product was more of an expression from various computer science communities of what a new research ethics framework may look like.

Challenges to the research ethics framework

Several controversies have arisen in the field of technical Internet research, though they have not been dealt with the same level of media attention nor wide scrutiny from scholars. The reason for this is unclear, though it could be related to the complex technical knowledge required to understand the research, the invisibility of the harms occurred, or simply because people did not feel affected by the consequences. This section does not describe case studies as these are integrated into the discussions of the concepts and the following case study chapters. This section does mirror the previous structure of the discussion on the challenges to existing research ethics frameworks by social science research conducted on the Internet. Each section illustrates how concepts of the research ethics framework developed for the biomedical sciences have evolved as they have been interpreted for technical Internet research.

Human subjects

In contrast to social science Internet research, technical Internet researchers have not traditionally researched human behaviour on the Internet. Computer science as a discipline has conducted relatively little research that involved human subjects directly (as defined by the Common Rule) (Buchanan et al., 2011). Increasingly, however, disciplines such as computer security and Internet measurement did begin

to involve human subjects, partly as research subjects or as a function of the ever-increasing socio-technical nature of the Internet. Dittrich and Kenneally note that unsuspecting Internet users may be implicated in this type of research as “owners or users of systems, networks, or devices that are subject to research” (Dittrich & Kenneally, 2012, p. 21). Furthermore, researchers frequently use datasets that have been leaked or obtained through other illicit means, and the data subject of the compromised datasets may not be aware that researchers are now using data related to him or her for research (Kenneally, 2015a; Thomas, Pastrana, Hutchings, Clayton, R., & Beresford, 2017). Research ethics review boards have not yet developed the capabilities to review this type of research. All too often, technical Internet research proposals are declared as being ‘exempt’ from full review because the members of the committees fail to identify the implicated human subjects that use these technologies (Buchanan et al., 2011; Partridge & Allman, 2016).

Computer scientists Craig Partridge and Mark Allman have observed the growing need for research ethics in their disciplines:

“Network measurement—because it is typically at arm’s length from humans—does not comfortably fit into the usual human-centered models for evaluating ethical research practices. Nonetheless, the network measurement community increasingly finds its work potentially affects humans’ well-being and itself poorly prepared to address the resulting ethical issues” (Partridge & Allman, 2016, p. 58).

They have called on computer scientists to identify when their research may be causing tangible harm to people, stating that it is the most important ethical issue for their discipline (Partridge & Allman, 2016). It is not always clear, however, when a risk of harm is created by technical Internet research or what the extent of the risk is

for an individual. Crandall et. al note that, due to the global nature of their work, risk is "highly contextually specific, and can vary depending on the location and background of the user running the measurement" and that "the legal and political environment from which it is run can significantly alter the level of relative risk" (Crandall et al., 2015, p. 1). As noted in previous sections, the lack of contextual knowledge complicates responsive actions.

Scholars have thus suggested alternative ways to interpret the binary nature of the definition of a human subject. For example, (Carpenter, K. & Dittrich, 2011) and (Aycock, Buchanan, Dexter, & Dittrich, 2011, p. 140) proposed a new concept of potentially "human-harming research," which would not focus solely not the implicated data subjects, but be more open to interpretation by ethics boards. This conception would include data collection from technical devices that are not necessarily interventions with living humans but may still cause potential harm to their owners or others. Furthermore, Buchanan has argued that the "distance principle" from social science Internet research ethics may also be relevant in determining whether a technical Internet researcher is aware of the extent to which humans are implicated in their work when they work on technical systems (Buchanan et al., 2011).

Minimal risk

Similar to the issues faced by social science researchers, the assessments of risk to data subjects are complicated by the global scale of technical Internet research. Risks of harm that are equivalent to the risks faced in everyday life in one place may in fact be detrimental to Internet users in another. For example, the risk of harm by exposing seemingly encrypted or obfuscated communications, may be slightly

embarrassing in a socially liberal country, but be cause for imprisonment (or worse) in an authoritarian regime. Accurately assessing the resulting harms and benefits for data subjects implicated in technical Internet research is therefore a challenge, especially for researchers trained mostly in a technical field (Buchanan, & Ess, 2008).

In an attempt to offer a framework for the risk of harm assessment, Partridge and Allman developed a spectrum based on which risks of harm can be assessed (Partridge, C. & Allman, 2016). While they argue that impact on people is merely a side effect of technical research—which focuses on the technical systems—they acknowledge that humans are increasingly implicated in the work of their community. Partridge and Allman differentiate between direct harm, indirect harm, potential harm, and the nature of the harm caused by Internet research. Direct harm can be inflicted on people or their systems by actions designed to target them as data subjects by the researcher. Indirect harms occur when the research targets technical systems for research, but humans are implicated in an unforeseen manner. Potential harm can occur when the actions of the researchers create the possibility of harm only when other events outside of their control occur. This is similar to the concept of downstream harms proposed by Buchanan, where either harms are experienced at a later moment or by persons who are not necessarily implicated in the research as data subjects but still experience harm as a result (Buchanan et al., 2011). The concepts of potential and downstream harms are relevant to understanding the impacts of research on society and are discussed in the case study chapters as well as the analytical chapter.

Autonomy

Literature addressing autonomy in technical Internet research is sparse. It focuses mostly on informed consent and has not (yet) addressed vulnerable persons or deception research. Crandall et al. highlighted a frequently recurring sentence included in technical Internet research papers: “[o]btaining prior informed consent for side channel measurements is not possible” (Crandall et al., 2015, p. 2). Indeed, it may be challenging to inform implicated but unsuspecting and anonymous Internet users and gain meaningful informed consent from them (Partridge, C. & Allman, 2016). Kenneally and Dittrich have uncovered three primary challenges to obtaining meaningful informed consent in Internet research, which highlights difficulties in assessing the relationship, context, and required responsiveness by researchers:

"[...] identifying who should consent to research; locating and communicating with the relevant parties; and explaining intangible and/or technical issues such as research risks in layman's terms" (Dittrich & Kenneally, 2012, p. 9)

Technical Internet researchers may prefer to ask their ethics review committees to waive the necessity for informed consent. A waiver can only be granted if the researcher can demonstrate that the research only poses a minimal risk of harm to data subjects and that it is impractical to obtain informed consent. This is a stricter requirement than the need to demonstrate that consent is merely difficult or inconvenient (Dittrich & Kenneally, 2012). Beyond the struggle for informed consent or a waiver, the concept of autonomy has not been given consideration in technical Internet research.

Privacy

Similarly to scholars in social science Internet research ethics, technical Internet researchers have struggled to operationalise concepts such as privacy and data protection into their systems or experimental design. This section interprets the three privacy-related issues from the experiences in social science Internet research ethics for the issues faced by technical Internet researchers. The public/private information distinction is exemplified by the data that is inevitably gathered when studying connected systems and devices. The blurring of anonymity and identifiability are considered through the lens of dichotomous legal and ethics frameworks. Finally, the reuse of data is considered from the perspective of longitudinal data collection and the sharing of research databases among researchers.

Public/private

The systems and devices studied in technical Internet research serve a variety of functions, which are mostly related to enabling the storage and transmission of information. To the technically proficient researcher, the information flows can be collected relatively easily, giving the appearance of this data qualifying for the exemption in research ethics of being “publicly available.” Further, experimentation on these systems may implicate unsuspecting Internet users or compromise their data—stored on devices or in transmission—for research purposes without the knowledge. Reasoning that accessible information is therefore publicly available exposes a deficiency in legal and ethical frameworks with regards to which information has been made intentionally publicly available, and which information is only accessible to those with the knowledge and tools to collect it (Kenneally, 2015a). Data collections over time and the sharing practices of these databases are

discussed in the subsection on reuse. This section briefly describes two methods of collection data: passive and active data collection.

In Internet measurement, passive data collections observe network data that is being communicated (Partridge, C. & Allman, 2016). The researcher would not inject data or initiate communications, though it may require devices to be installed in a network (Chaudet, Fleury, E., Lassous, I. G., Rivano, H., & Voge, M.-E., 2005). The focus could be a specific end-user device or a server that interconnects many devices. For example, researchers may directly access log files of websites or from the Domain Name System (DNS) to discern how, when, and possibly who accessed information online (Khan, M. T. & Kanich, 2015). In some instances, for example at telecommunication providers, it is also possible to set up so-called middle box devices that decrypt and read information flows (See for example (Naylor, et al., 2015)). Such passive data collections can be considered as conducting surveillance on unsuspecting Internet users, however. To avoid this, researchers have developed applications that are installed or accessed by their research subjects on their devices. For example, the Netalyzr project provided an applet which is accessible through web browsers (Kreibich, Weaver, Nechaev, B., & Paxson, 2010) and the Haystack project developed a dedicated mobile phone application (Razaghpanah, et al., 2015).

In addition to Internet measurement, researchers can find or encounter data stored online that may be beneficial for their study. Researchers have used (and published research papers based on) data from other research projects or databases which have been made access through illicit means. For example, it is relatively easy to acquire databases that contain leaked username and password combinations for online services, hacked email databases, information about device vulnerabilities, or even

health and financial records. Scholars have stated that it is currently not clear what the responsibility of the researcher is when they obtain existing data that is either the result of malicious and illicit activity, or when it was poorly secured from public browsing (Dittrich & Kenneally, 2012; Kenneally, 2015a; Thomas et al., 2017). Clarity about the level of required responsiveness would therefore benefit from better guidelines for interpretation.

As opposed to collected data that can be observed or found, researchers can also actively inject data or perform other deliberate actions to acquire data. In technical Internet research, this can take many forms. For example, a researcher may probe systems by sending requests to measure a network topology (Willinger, W. & Roughan, 2013), employ a botnet which is distributed across many devices to initiate measurements in new networks (Anonymous, 2012), or even initiating a web-crawler to collect data from the worldwide web (Thelwall, 2001; Thelwall & Stuart, 2006). Further, information and computer security researchers may actively intrude in systems to assess their security vulnerabilities (Matwyshyn, Cui, Keromytis, A. D., & Stolfo, 2010).

Whether data is collected through passive observations or active and deliberate intrusion, there is typically a likelihood that data is collected about individual and potentially identifiable persons. These persons may be unsuspecting Internet users, who are implicated in a research study without their knowledge nor their consent. The data may be available for collection if one has the technical knowledge and tools required to access it, though such data will likely be in tension with the Internet users' perceptions of privacy and their data protection rights.

Anonymity and identifiability

The impact of re-identification techniques on the binary concepts of anonymity and identifiability has been studied extensively by the technical Internet research community. Previously ‘anonymised’ datasets that were published for public use have been repeatedly re-identified (See for example the body of work by Arvind Narayanan (Narayanan, A. & Shmatikov, 2006; Narayanan, A. & Shmatikov, 2008; Narayanan, A. & Shmatikov, 2009; Narayanan, A. & Shmatikov, 2010; Narayanan, A. & Felten, 2014)). Other techniques allow inferences to be drawn about individuals based on seemingly non-identifying data types. For example, geographic human mobility traces captured from mobile phones have been found to be highly unique. Only four location data points over the course of one day were necessary to identify 95% of individuals in a dataset (Montjoye, Hidalgo, C. A., & Verleysen..., 2013). The settings, configurations, and combination of plug-ins of certain applications—such as a web browser—can also be used to distinguish individuals and identify them with adequate auxiliary data (Doty, 2016). Re-identification techniques typically employ other available auxiliary data, which may be linked through unique identifiers and may provide enough contextual information to re-identify individuals (Buchanan, & Ess, 2008; Dittrich & Kenneally, 2012).

Enriched databases (i.e., datasets that have been fused with other datasets with more information about data subjects) may also enable the re-identification of data subjects, thus breaching any previous assurances of anonymity or confidentiality (A29WP, 2014; Narayanan, A. & Shmatikov, 2009; Ohm, 2009; Schwartz, P. M. & Solove, 2011; Solove, 2012). Furthermore, enriched databases may magnify or generate new sensitivities that exist in the primary dataset (Kenneally, 2015b). Enrichment can thus expose previously unknown—or purposely separated—

information about data subjects (Crawford, K. & Schultz, 2013; Schwartz, P. M. & Solove, 2011; Solove, 2012).

In response, legal scholar Paul Ohm wrote a paper titled 'Broken promises of privacy: Responding to the surprising failure of anonymisation' in which he argues that full de-identification is not possible without significant loss of data utility. The dichotomous nature of the law fails to take into account researchers' technical ability to re-identify previously anonymised or de-identified databases (Ohm, 2009). Statutory exemptions in existing data protection laws or research ethics frameworks are therefore no longer suitable standards to be applied in technical Internet research.

Scholars have suggested that the identifiability of research subjects should be considered on a continuum, too. For example, Swartz and Solove propose a new approach to the existing legal approach based on the 'risk of identification', offering new categories to be considered (Schwartz, P. M. & Solove, 2011). Similarly, El Emam has suggested an 'identifiability continuum', which bases a five-level identifiability model on both the level of aggregation and the risk of re-identification according to the capacity and capability of potentially interested threat actors, such as prosecutors, journalists, or marketers, amongst others (El Emam, 2010). Responsibilities for researchers would therefore depend on the level of (re-)identifiability of their data subjects.

Reuse

Similarly to their colleagues in the social sciences, technical Internet researchers can reuse existing data in a variety of ways. In fact, researchers are encouraged or even

required to share and reuse their datasets to further efforts to reproduce academic work (NSF, 2011; Wood, et al., 2010). Computer scientists have shared a large amount of research datasets, collected and managed in repositories such as Crawdad,⁵² the Cooperative Association for Internet Data Analysis (CAIDA),⁵³ and the Measurement Lab.⁵⁴ Reusing such datasets allows researchers to circumvent the research ethics review—at least in the US—as reuse would fall under the exemption “involving the collection or study of existing data” in §46.101 (b) of the Common Rule.

Technical Internet researchers have noted some issues with applying a blanket exception for existing data. First, the data may contain or reveal sensitive aspects of people’s lives. Individual Internet users may be identifiable, as IP addresses are typically stored to connect databases over time. Second, as discussed in the section on public and private data, reusing data which was collected or made accessible in illicit ways would also be considered a simple reuse of that data, exempting researchers from ethics review requirements. Partridge and Allman argue that such data can provide meaningful insights, which could be viewed as “using the data as making the best of a bad situation.” However, they suggest that the use of such data may also justifiably be considered as a “moral hazard” (Partridge, C. & Allman, 2016, p. 62). The case studies in this thesis shed some more light on these issues.

⁵² Available at <<https://crawdad.org/>>.

⁵³ Available at <<http://www.caida.org/data/overview/>>.

⁵⁴ Available at <<https://www.measurementlab.net/data/>>.

Towards a New Internet Research Ethics?

As Internet research ethics has grown as a field tangential to social science Internet research, it became increasingly clear that existing guidelines for ethical research on human subjects in biomedical or social sciences could not always be usefully applied to Internet projects (Eysenbach, G. & Till, 2001). In response to scandals in Internet research, such as those described above, some scholars even called into question the appropriateness of existing research ethics frameworks for emerging forms of online research and called for experimentation with completely new models of governance for Internet research scholarship (Bassett, E. H. & O'Riordan, 2002; Buchanan et al., 2011). In fact, as early as 1995, Lynne Schrum noted the need to reinterpret the concepts of research ethics in her paper titled 'Framing the Debate: Ethical research in the Information Age':

"Those who use on-line communications have moved far beyond the expectations of the first networkers [...]" and that "[t]he concept of ethical research requires new exploration and definition" (Schrum, 1995, p. 311)

Indeed, the collection of data about people via digital networks or assessment of technical Internet research projects that also implicate humans and their devices gave rise to new problems, which did not fit smoothly into the existing regulatory language of research ethics (Aycok, Buchanan, Dexter, & Dittrich, 2011; Buchanan et al., 2011). The human subject model of research ethics focuses in large part on the collection, processing, and dissemination of research data as it relates to human subjects. The changing characteristics of information, enabled by the dynamic informational environment created by the Internet, affect the application of research ethics laws and frameworks that are based on the assumptions of a static informational environment, similar to other information laws.

Scholars have begun to address the effect of the new informational environment on research ethics frameworks. William H. Dutton and Tina Piper note that "e-research does pose some novel issues and conflict across competing values" for a traditional research ethics approach to be sufficient for diligent oversight (Dutton & Piper, 2010, p. 235). Michael Zimmer and Katharina Kinder-Kurlanda observe that the Internet and related technologies have opened up "vast new means for communication, socialization, expression, and collaboration" and wonder to what extent "the Internet [has] changed the foundation from which considerations of traditional research ethics begin" (p. xxix). For example, they specifically note that it is increasingly becoming the norm for Internet researchers to aggregate data from disparate sources—whatever auxiliary data may be useful and available—thereby making use of the characteristics of the environment created by the Internet in a way that would have been much harder or even impossible without the Internet (Zimmer, & Kinder-Kurlanda, 2017).

It was further noted that researchers themselves may not be fully aware that their projects have entered the domain of 'human subject research', whereby ethical responsibilities arise (Buchanan et al., 2011). Overly simplistic applications of spatial metaphors, such as references to the Internet as a 'cyberspace' to link the network to physical phenomena have been considered to be misleading for the useful application of ethics frameworks to Internet research (Bassett, E. H. & O'Riordan, 2002). Furthermore, existing frameworks may be unsuitable for "[...] the need to accommodate the different cultural, legal and social contexts of the people around the world who participate in such studies" (Eynon et al., 2009, p. 197). Thus, Keller and Lee note that "conducting human participant research using the Internet can be

fraught with ethical landmines, some of which are yet to be discovered” (Keller, H. E. & Lee, 2003, p. 218).

Furthermore, technically proficient researchers can use information technologies to—wittingly or unwittingly—subvert some of the basic principles and concepts of research ethics. For example, informed consent statements can be presented in ways that make it easy for the Internet users to accept them, encouraging research participants to forego reading relevant information. These methods circumvents the assumption that researchers and data subjects are in the same location, are able to speak about the research and its impact, and come to a mutual understanding about what they consent to. Researchers can also covertly make use of other peoples' devices in a variety and somewhat deceptive or subversive ways, whereby their owners may not be aware that their devices are being used in a research project.

A culture of reviewing research projects through the lens of human participants may be lacking in technical disciplines, however. Disciplines such a mathematics, computer science, network engineering, or even statistics do not historically consider themselves to be conducting human subject research as in the biomedical tradition (Metcalf, J. & Crawford, 2016). Humans have not typically been the objects nor the subjects of technical R&D projects. Therefore, considerations of the interests of humans implicated in a technical study can, and often do, fail to be taken into account in interpretations of human subject research ethics laws and guidelines.

Need for new frameworks challenged

Other scholars have noted, however, that established principles of research ethics should not be cast aside by blind enthusiasm for the new medium. For example,

Alun Jones has stressed that researchers should not lose sight of their responsibilities in this new field and understand how online research activities violate existing professional codes of ethics and research ethics frameworks (Jones, 1994). Similarly, Sharon Boehlefeld stated in 1996 that the goal of research ethics still remains to do good and avoid doing harm. She has suggested that the term ‘cyberspace’—as it is still sometimes used—may indeed mystify the medium through which research is conducted, but that the research practices should not be assessed differently than traditional forms of research. She explains that the enthusiasm of early Internet researchers did indeed lead to the fundamental research ethics principles being forgotten:

“When a new phenomenon crops up, like the Internet, questions arise that seem so important or exciting that researchers might, in their enthusiasm, focus more on those questions than on the fact that there are people on the other side of their data. When that happens, thinking about ethics may help put the human side of the phenomenon in the forefront again.” (Boehlefeld, 1996, p. 141)

A range of scholars have agreed with this interpretation. Scholars like Elgesem state that Internet research does not raise issues that are different from other types of research in the humanities or social sciences (Elgesem, 2002). Eynon et. al warn that considerations must be sensitive to the offline context within which data subjects are nested. This may require researchers to interpret and operationalise laws, values, and their tensions in a variety of context-dependent and culturally-specific ways (Eynon et al., 2009). Despite recent challenges to research ethics frameworks, scholars like Ess have stated that ethics guidelines related to the context of the Internet are not special enough to warrant a new framework.⁵⁵ These scholars reject

⁵⁵ This document can now only be found through the Internet Archive:
<https://web.archive.org/web/20021215165316/http://www.nyu.edu/projects/nissenbaum/ethics_ess.html>.

the proposal of a new framework of research ethics for Internet research and call for further interpretation of the human subject model.

Experimentation

Taking up the challenge of experimentation with new research ethics review frameworks, David Berry described an open-source approach to Internet research ethics, suggesting a participatory model between stakeholders to take into account the interests of researchers, policymakers, enforcers, and research participants (Berry, 2004). A similar approach called Responsible Research and Innovation (RRI) is being developed within the European Union (Commission, 2012). Scholars note that such participatory models for research and innovation on emerging technologies such as the Internet inherently involve political discussions about "[...] power, democracy, and equity" (Owen et al., 2013, p. 37. See also Grunwald, 2011; Stahl, 2013; van den Hoven, 2013). These topics have not been addressed in existing research ethics frameworks.

Conceptual framework (1)

This section presents a summary of the chapter in four tables. Each table contains the newly uncovered interpretations of a research ethics concept (i.e. human subjects, minimal risk of harm, autonomy, and privacy) developed in this chapter, through the lens of the ethics of care research themes. The vertical columns contain the high-level research ethics themes (ie. relational approach to morality, contextually-oriented deliberation, responsiveness, and emotions/attitudes. The interpretations of the concept are further divided into the aspects of the research theme as outlined in

Chapter 2. Together, these four tables form the first iteration of the conceptual framework of this thesis.

Human subjects	
Relational approach to morality	Data subjects: • Technological mediation ◦ Relationship established through the use of networked technology, ◦ Enrolment into relationship may not be voluntary, ■ Consent technically not necessary for interaction or intervention • Distance principle ◦ Disintermediated relationship through affordances of networked technology, ◦ Increased experiential distance of researcher ◦ Researchers and data subjects not easily identifiable ◦ Data subject is <i>distant other</i>. Society: • Hierarchical societal position of researcher ◦ Academics enjoy a powerful position, ◦ Technologists shape human communications and interaction, • Random people can be implicated in an Internet research study by virtue of using the Internet. • Power relationship to data subject, ◦ Potentially oppressive relationship, ◦ Data subject is vulnerable to the decisions made by researchers, if lacking decision capacity. Research profession: N/A
Contextually-oriented deliberation	Research context: • Assessing the implications of a disintermediated relationship is based on a lack of knowledge. Data subject context: • Difficulty or impossibility to know local context of data subjects, • Replace dichotomous human subject concept with continuum based on a variable such as identifiability or potential risk of harm.
Responsiveness	Identify: • Human ownership or use of implicated devices, Understand: • Data collection through intervention or interaction with devices may implicate humans, Interpret: • Sensitivity to the offline context within which the data subject is likely nested.
Emotions/attitudes	Positive: • Beneficial research, ◦ By benign researchers. • Possible to understand and meet the expectations of Internet users. Negative: • Humans may be used as a means to an end, ◦ If local context is not identified and understood, ◦ And researcher is not responsive to their specific needs. • Research can be considered as a form of surveillance.

Minimal risk of harm	
Relational approach to morality	Data subjects: N/A Society: N/A Research profession: N/A
Contextually-oriented deliberation	Research context: • A single 'daily life' standard does not exist, ◦ Variables in the typical calculation of the magnitude of risk multiplied by its probability are too complex and vague to be meaningful, • Researcher may not be able to assess whether their actions constitute a minimal risk of harm. Data subject context: • Harms are highly contextual, • Data subjects can be situated in wide-ranging contexts with undefined sociopolitical parameters, • Research may create new risks of harm in the data subject's context: ◦ By disregarding local (informational) sensitivities, ◦ By being insensitive or ignorant to existing power relations, ◦ By violating local laws.
Responsiveness	Identify: • Human ownership or use of implicated devices, • Context-specific harms, • Unsuspecting data subject expectation and perception should be leading: ◦ Researchers to seek out rules, etiquette, guidelines that govern the virtual space or community. Understand: • Researcher may not be able to understand or generalise the data subjects' needs, interests, and the context-specific risks of harm, • Harms may only become apparent after the data collection or publication of the research. Interpret: • Which standard is the researcher imposing on all data subjects? • Interpret minimal risk standard on a continuum, instead of the static and universal 'harm [...] no greater than those ordinarily encountered in daily life' standard, ◦ Differentiate between: ■ Direct harm, ■ Indirect harm, ■ Potential harm.
Emotions/attitudes	Positive: • Benevolent intentions Negative: • Possibility of conducting deception research without protocol, • Exploitation of vulnerable persons, ◦ Implicated Internet users are not aware of the research, ◦ Implicated Internet users cannot exercise their rights.

Autonomy	
Relational approach to morality	Data subjects: • Unknown or misleading identities, ◦ Researchers, ◦ Data subjects. • Status of relationship is unclear if identities are unknown, Society: • Researchers can enrol Internet users seemingly at random ◦ Diminishing the ability of all Internet users to choose to take part in research, • Lacking clarity whether informed consent establishes autonomy or a relationship meaningfully in research at all, • People's interactions with those who have consented, can be implicated, • Ability by researchers to alter psychological state of Internet users • Ability by researchers to change Internet users' environments. Research profession: N/A
Contextually-oriented deliberation	Research context: • Can researchers infer the intentions and expectations of users to provide existing data? • The researcher may be required to make an effort to make themselves known to data subjects, • Difficulty or impossibility to identify vulnerable persons or groups within the research population, • Communication about foreseeable risks is not feasible without a contextual analysis, • Notification of rights (e.g. to withdraw) or debrief after deception may not be possible. Data subject context: • Difficult or impossible to understand data subject context.
Responsiveness	Identify: • Are data subjects in a position to give their informed consent? • Uncertainty whether research subjects understand what they consent to, • Competence to provide informed consent may vary, • Vulnerability status (decisional capacity) of research subjects, ◦ If lacking decisional capacity. Understand: N/A Interpret: • Consent via proxy: is the person giving consent (e.g. from website or platform owner on behalf of their users) legitimised to do so?
Emotions/attitudes	Positive: • Respecting autonomy through engagement with data subjects, Negative: • Potential violation of the freedom from being subjected to research, • Outside interaction with a person's device may be a violation of rights, • Violations of dignity.

Privacy	
Relational approach to morality	Data subjects: - Relationship to researcher established through networked technology connectivity, - The identity of Internet users is intertwined with their devices due to increasing sociotechnical convergence, - Re-identification in reused datasets may establish a relationship to previously anonymous data subjects, - Those with technical capability to collect data, establish the relationship. Society: - Researchers can in theory collect research data about any Internet user, - Without the Internet users need for knowledge or consent for data collection. - The power to collect data establishes a power hierarchy, - A form of surveillance. Research profession: N/A
Contextually-oriented deliberation	Research context: - Distinction between public and private domains are blurred. Data subject context: - For passive data collection, the reuse of data in research context may not be perceived as appropriate.
Responsiveness	Identify: - User expectations and perceptions of privacy in their online context, - For passive data collection, identify the purpose for which data is originally shared, - Potential for re-identification after anonymization of dataset. Understand: - Sensitivity of data for data subjects, - Obligation to exercise special caution, - When data subjects do not know they are participating or data is collected about them, - Lines between data collection, illicit intrusion into devices, and deception, - For passive data collection from an illicit source, the responsibility for researchers may be unclear. Interpret: - Special caution when collecting data on unsuspecting data subjects, - Tension: Data utility vs. purpose limitation, - Determine level of privacy to be respected based on: - Group accessibility, - Perceived privacy of data subjects. - Offline analogies should not replace assessment of user expectations.
Emotions/attitudes	Positive: - Control of data, - To maintain dignity. Negative: - Potential for surveillance, - Loss of privacy, - Reported feeling of being dissected like laboratory animals, - Internet users become a means to an end for researchers. - Feeling of being violated.

(4) Case Study 1: Password Dumps in Research

Introduction

The first case study chapter analyses a study titled *The Next Domino to Fall: Empirical Analysis of User Passwords across Online Services*, authored by Chun Wang, Steve T.K. Jan, Hang Hu, Gang Wang at the Department of Computer Science of the Virginia Polytechnic Institute and State University (popularly known as Virginia Tech). The paper was published as part of the proceedings of the ACM Conference on Data and Application Security and Privacy, held Tempe (Arizona) in March 2018. A previous iteration of the paper titled *Empirical Analysis of Password Reuse and Modification across Online Service* came to my attention during a literature research of password dumps. This version was uploaded to the popular website for computer science papers www.arxiv.org on June 6th 2017.

This chapter is divided into three parts. The first section provides a description of the case study paper and situates the study within its wider research domains. The second section analyses the research paper through the lens of the applicable research ethics framework, which is the US Common Rule that was applicable at the time the research team would have gone through the application procedure (before the revision of 2018 came into effect).⁵⁶ This section identifies the areas where the research ethics framework leaves ambiguity or may be insufficient. The third section

⁵⁶ The Common Rule has been amended and revised, though it has not yet become fully effective due to ongoing delays. See <<https://www.federalregister.gov/documents/2018/06/19/2018-13187/federal-policy-for-the-protection-of-human-subjects-six-month-delay-of-the-general-compliance-date>>.

assesses to what extent the ambiguities that surface in the analysis can be addressed by the existing scholarly debates covered in the conceptual framework, or to what extent new issues arise. This section iterates on the conceptual framework, which informs the analysis of the next case study.

Contextualising the case study paper

The paper by Wang et. al is exemplary for computer science research that uses data from illicit sources, such as password dumps. This section begins by situating the case study paper within its research domain. First, the academic debate around computer science research that makes use of illicit data, or data from illicit sources, is discussed. Then, the next section reviews computer science research and the ongoing discussions about the appropriateness and desirability of using password dumps in computer science research, which has a long tradition. Finally, the last subsection describes the case study paper and highlights how statements in the paper connect to the existing literature.

Research on illicit data

Similar to other disciplines of research, computer scientists may generate, collect, and re-use data for their research. Some of this data can be obtained from dubious sources or by questionable means. Depending on the sources and methodologies of collection, some datasets may be considered illicit or in violation of widely held norms. For example, the computer science community has in the past considered datasets to be illicit when they are obtained by unlawful intrusion into systems, or by

reusing leaked databases (such as passwords, classified materials, or financial and medical data). Daniel R. Thomas et al. describe illicit data as:

"[...] data collected as a result of (i) exploitation of a vulnerability in a computer system; (ii) unintended disclosure by the data owner; or (iii) unauthorized leak by someone with access" (Thomas et al., 2017, p. 445).

In Thomas et al.'s first category, the researcher creates primary research data, while in the second and third, existing data is collected for another purpose and reused in research. Password dumps fall into the second category, though it is relevant to explore both categories to understand the differences in practice and their meaning with regards to the research ethics framework.

(i) Exploitation of a vulnerability in a computer system (Primary data)

Internet researchers can actively gather primary research data through illicit means by knowingly and deliberately initiating a data collection that violates context-relevant norms or applicable laws. This includes scraping websites in violation with terms of service or data protection laws, unlawful intrusion into information systems, or otherwise exploiting information security vulnerabilities by installing additional software on end-user devices (e.g. malware). For example, as described in previous chapters, a paper published by anonymous researchers describes how they created a topological map of the Internet through a comprehensive IPv4 survey of the Internet (Anonymous, 2012). The researchers created a botnet that infiltrated systems and subsequently spread the self-replicating code to other systems within the networks of the infected system. In total, they intruded into 420,000 information systems by exploiting default credentials of an entry to the Linux operating system upon which these systems operated. The collected data was vast and the inferences useful, though the botnet infiltrated systems connected to important physical

systems, such as “[...] industrial control systems, and physical door security systems” (Dittrich et al., 2012, p. 2-3). This method of collecting data—through malware which exploits a vulnerability and intrudes into systems—was thus considered to be in violation of ethical standards by several researchers in the computer science community (Dittrich et al., 2012).⁵⁷ However, and interestingly, the computer science community as a whole would not unanimously condemn the further reuse of the collected data to be in violation of ethical norms.⁵⁸

(ii) Unintended disclosure by the data owner and or (iii) unauthorised leak by someone with access (Secondary data)

Datasets can be disclosed intentionally or by human error, using the tools and methodologies described in the first category above. The disclosed, leaked, or hacked datasets can typically be accessed freely online, or for an incidental charge. For example, several leaked databases are advertised on the www.pastebin.com website.⁵⁹ Other examples include leaked classified information available on platforms such as Wikileaks⁶⁰ and financial data such as the Mossack Fonseca leak also called the 'Panama papers'.⁶¹ Researchers can technically reuse these databases for their research, though research ethics frameworks may impose limits.

The research ethics issues concerning the use of illicit secondary data were recently brought to light again when in 2015 the Patreon website—a crowdfunding website for artists of various sorts—was hacked and all their data and source code was

⁵⁷ See <<https://www.caida.org/research/security/carna/>>.

⁵⁸ See <<https://www.wired.com/2013/05/internet-census/>>.

⁵⁹ See for example the leaked Twitter password dump available for 0.5 Bitcoin: <<https://pastebin.com/n9phcmx4>>.

⁶⁰ See <<https://wikileaks.org/>>.

⁶¹ Accessible at <<https://offshoreleaks.icij.org/pages/database>>.

leaked.⁶² This database contains the source code, data of all projects, and data about users, such as names, private messages, including their email addresses as identifiers. Two researchers, Nathaniel Poor and Roei Davidson, were already conducting Internet research on the Patreon website prior to the breach. They initially did not have access to all the data they needed for their research, and this leaked database provided all the missing data points.

While the researchers considered the leak as a "gift" to their research project, they were also aware that the use of this data may violate existing and applicable research ethics frameworks. They asked the Association of Internet Researchers mailing list for advice and guidance regarding the research ethics of using data obtained by illicit means (ie. hacking) for their research project.⁶³ Based on the responses, Poor and Davidson published a paper that discusses the reasons why they decided to refrain from using the hacked data for their project. In short, considerations of user's expectation of data use, potential violations of privacy by the researchers, the absence of informed consent, legitimising criminal activity, and other potential means to collect the necessary for their research were the factors that outweighed their reasons to use the data (Poor & Davidson, 2016).

Password 'dumps'

The term "password dump" is the colloquial term for an involuntarily disclosed database that contains passwords with their corresponding usernames, which are credentials to access an online service. Passwords can be disclosed both by hacking into a service (exploiting an SQL injection vulnerability technique appears to be

⁶² Accessible at <<https://raidforums.com/Thread-Patreon-Database-Leaked-Download>>.

⁶³ The email is available at <<http://listserv.aoir.org/htdig.cgi/air-l-aoir.org/2015-October/251741.html>>.

common (Weir, Aggarwal, Collins, & Stern, 2010)), and by unintentional or unauthorised leaks. Several prominent examples exist, such as the Rockyou password dump breached in 2009 containing 32 million individual data points,⁶⁴ to the Myspace password dump made available in 2016 containing 427 million passwords.⁶⁵

Research papers typically refer to password dumps being available on websites such as www.pastebin.com and www.pastie.com, as well as forums or marketplaces that are accessible through specialised software. Some password dumps are available for free and accessible by clicking on a link or just searching for them with common Internet search engines. Others will only be disclosed after requesting them from individuals, which may require the requester to pay a fee. There are several research communities that focus specifically on password research, which host annual conferences.⁶⁶ The next section describes the general trends in password research, their aims and purpose, and the typical justifications researchers put forward for conducting research on this type of illicit data.

Aims and Purpose of password research

The dominant security mechanism for the identification of individual Internet services is the combination of a unique username and a memorised password (Bonneau, 2012). While passwords may not be an absolutely reliable mechanism to authenticate individuals (e.g. passwords may be forgotten, shared or leaked, and

⁶⁴ See <<https://www.helpnetsecurity.com/2009/12/14/serious-sql-flaw-could-have-compromised-millions-of-rockyoucom-users/>>.

⁶⁵ See <<https://mashable.com/2016/07/01/myspace-password-database/#igHeEovcaigt>>.

⁶⁶ See for example Cambridge, UK in 2005:<<https://www.cl.cam.ac.uk/events/passwords2015/>>, Bochum, Germany in 2016: <<https://passwords2016.rub.de/>>, Stockholm, Sweden in 2018: <<https://passwordscon.org/2018/07/passwordscon-november-19-20-2018-in-stockholm/>>, a conference in 2017 was to be held in Las Vegas, but cancelled due to the travel ban imposed by the US Government.

guessed or inferred by third parties), their ongoing widespread use today indicates that they have stood the test of time. Researchers who analyse passwords and the security mechanism within which they are used, are typically interested to discover and mitigate their inherent flaws. This requires an understanding of how passwords are created, reused across services, how the mechanism may be compromised, how password strength meters (PSMs) may be optimised, or other policies that set a minimum threshold for the complexity and thus the strength of a password against known attacks.

Password composition and reuse studies

With the aim of improving the security and authentication mechanism provided by passwords, Hanamsagar states that “[i]t is critical to study both how users choose and reuse passwords, and the causes that lead users to adopt unsafe practices” (Hanamsagar, Woo, Kanich, & Mirkovic, 2018, p. 1). Studies on passwords were conducted before large-scale password dumps became available (see for example (Florencio & Herley, 2007). The availability of password dumps negated the need to set up elaborate experiments to gain sufficient data about passwords to make generalizable inferences. For example, Das et al. “[...] study several hundred thousand leaked passwords from eleven web sites and conduct a user survey on password reuse” to find that an estimated 43-51% of Internet users reuse the same password for multiple services (Das, Bonneau, Caesar, Borisov, & Wang, 2014, p. 1). To underpin such measurements with sound scientific methods, Bonneau et al. used password dumps to “[...] formalize improved metrics for evaluating the guessing difficulty of a skewed distribution of secrets, such as passwords [...]” (Bonneau, 2012, p. 1). From a theoretical point of view, Malone and Maher state that they “[...] collected sets of passwords belonging to sites which were compromised

and the lists of passwords subsequently publicly leaked” with the aim to “[...] investigate if Zipf’s law is a good description of the frequency with which passwords are chosen”⁶⁷ (Malone & Maher, 2012, p. 302). On a more practical level, Li et al. use password dumps to investigate the extent to which Internet users include personal information, such as names or birthdays, in their passwords, which may allow potential attackers to guess the passwords more easily (Li, Wang, & Sun, 2016).

Understanding password attacks

Researchers use password dumps to learn how the password security and authentication mechanism can be attacked. For example, Ur et al. have made use of password dumps to assess how this data can be leveraged by attackers to guess passwords based on popular known techniques (Ur et al., 2015). Similarly, Das et al. used hundreds of thousands of leaked passwords to understand how attackers can exploit Internet user’s reuse of passwords across services (Das, Bonneau, Caesar, Borisov, & Wang, 2014). Veras et al. show how attackers may use natural language processing techniques on password dumps to more efficiently inform their dictionary-based attacks (Veras, Collins, & Thorpe, 2014). Similarly, though on a much smaller password sample, Narayanan and Shmatikov show how Markov models (from probability theory) can be used to dramatically reduce the size of probable passwords that an attacker would attempt, thereby improving the guessability of some password combinations (Narayanan & Shmatikov, 2005).

Password strength

An understanding of how attackers might attempt to guess passwords allows researchers to study how passwords—and the mechanisms that *measure their strength*

⁶⁷ Zipf’s law states, in short, that given a large sample of words used, the frequency of any word is inversely proportional to its rank in the frequency table.

against common attackers—can be optimised for security. Password dumps provide an opportunity for researchers to quantitatively understand the threats to passwords and to users once their passwords have been leaked, as well as develop countermeasures for common attacks (Ji, Yang, Das, Hu, & Beyah, 2017). For example, Weir et al. make use of password dumps to understand the extent that “word-mangling” (ie. the changing characters in a password to resembling symbols, e.g. changing ‘password’ to ‘p@ssword’) improves the strength of passwords (Weir, Aggarwal, De Medeiros, & Glodek, 2009). Ma et al. find that password composition policies can benefit from the statistical language modeling literature, which had previously been unexplored (Ma, Yang, Luo, & Li, 2014).

Further, password dumps allow researchers to assess the usefulness of PSMs, as well as the validity of assumptions and evidence on which they have been developed. For example, Komanduri et al. identify commonly held beliefs about password composition are inaccurate based on samples of the RockYou dataset, though have been included in policies underlying PSMs (Komanduri et al., 2011). Other scholars have published similar findings and conclude that many PSMs fail to serve their purpose. Carnavalet and Mannan demonstrate inconsistent recommendations of a variety of PSMs, when presented with the same passwords. Objectively weak passwords are sometimes labelled as ‘strong’ or even ‘very strong’ by some PSMs (De Carnavalet & Mannan, 2014).

Justifications

A review of the corpus of password research suggests there are four main justifications for using password dumps as research data. Researchers rely on (1) the (now) public nature of the data leaked data, (2) stating that other researchers have

set the precedent to make use of password dumps for research purposes, (3) that use of password dumps entails low or minimal risk for the data subjects, and (4) that their benevolent intentions ensure great benefit, as their research output can help to hinder malevolent actors to use the password dumps for ill-intentioned purposes in future. The four justifications are described in more detail below.

1) Publicly available data:

Several papers include a variation of a sentence such as "we only analyse datasets that are already publicly available." This is a mere technical view of data that is, indeed, no longer secured due to hack or a leak, but disregards the legality or violation of ethical norms of using such data. However, the researchers may feel justified due to statements such as the "[i]nvasion of privacy liability for use of online data for research is not triggered since the researcher is not the actor who originally exposed the user's private information online" (Kenneally, 2015b, p. 4). This line of reasoning seems to give researchers a self-serving free pass to use leaked data, such as password dumps, as long as they did not cause the leak in the first place.

2) Precedent has been set:

Password dumps have been extensively used in computer science research for benevolent ends. Researchers in more recent papers typically refer to existing peer-reviewed literature and conference proceedings as a justification that using password dumps must have therefore passed ethical scrutiny elsewhere. While there may be some merit to an argument like this, there may also have been a lacking or uninformed ethical scrutiny of research using password dumps in the past, which allowed this type of research to pass into publication in the first place. It also raises

the question whether a pervasive violation of research ethics lessens the magnitude of a further violation.

3) Low risk to users:

A common argument in the password dump literature is to claim a low additional or minimal risk to data subjects, because their passwords have already been exposed.

Researchers argue that this type of research advances science and allows researchers to help develop better security mechanisms that protect users in the longer term.

This argument justifies the means of using data from an illicit source by potential, speculative ends from the research.

4) Benevolent researcher:

Finally, most researchers claim that their own good intentions justify their use of password dumps for data. They warn that the datasets are also available to potential malevolent attackers, so obstructing this type of research may give those malevolent attackers an advantage. The use of password dumps allows researchers to assess and develop defensive techniques based on the same datasets that malevolent attackers may have.

Concerns and criticisms

Responses to these justifications can be found in the scientific and ethics literature about the moral nature of using password dumps for research.

1) Publicly available data:

In the proceedings of a panel organised in 2012 at the Financial Cryptography and Data Security workshops on Bonaire titled *It's Not Stealing If You Need It: A Panel on*

the Ethics of Performing Research Using Public Data of Illicit Origin, Stuart Schechter argues that researchers should not consider a compliance test to be equivalent to ethical consideration. While research ethics provides an exemption for 'publicly available data,' this does not mean that a password dump is "fair game." Schechter suggests understanding the further use of data from the perspective of the user: would they have intended to have their data available for research? (Egelman, Bonneau, Chiasson, Dittrich, & Schechter, 2012). The scholarly discourse in Internet research ethics has previously addressed similar situations. For example, Buchanan and Ess discuss the use of archival data. They argue that researchers and ethics review boards should consider the ramifications of using pre-existing data from archives, because existing data about persons can come back to haunt them if it is brought back to light through research. The use of publicly available data may therefore violate the spirit of a consensual research relationship (Buchanan, & Ess, 2008). Similarly, Thomas et al. argue that data of illicit origin is unlikely to have been intended for research purposes or public exposure. They suggest that the data must be anonymised or data subjects' privacy preserved in other ways, if this data is to be used for research (Thomas et al., 2017). Finally, while the databases—or password dumps—may now indeed be technically publicly available as a result of the leak, it was the intention of Internet users and the service to keep this data secret. This data may be considered personal data (under European law) or personally identifiable information (in US law), so therefore its processing for research purposes may be a violation of data protection and privacy laws. This legal restriction should be weighed when discussing the availability of data.

2) Precedent has been set:

Allman and Partridge raise the question whether the research use of data from illicit origins legitimises or condones the act of collecting the data in the first place (Partridge, C. & Allman, 2016). An ethics board would likely consider a research study to be in violation with research ethics frameworks and even unlawful if the researchers had themselves actively compromised a set of privately-owned devices to obtain password dumps. When researchers gain professionally (e.g. publications, promotions) from research using data gained through an illicit action performed by another person, are they incentivising others to use this type of data, too? And, by extension, do computer science communities and their research institutions therefore condone and legitimise the use of password dumps and other data from illicit sources?

3) Low risk to users:

The proceedings of the panel titled *It's Not Stealing If You Need It: A Panel on the Ethics of Performing Research Using Public Data of Illicit Origin* is also illuminating with regards to the minimal risk requirement, or the balancing of risks and harms. Joe Bonneau uses a utilitarian argument that the scientific potential of illicit datasets is "too large to ignore." He suggests applying lessons learned in security vulnerability research to the use of password dumps, such as applying a 'do no harm' principle, applying the strict rules of responsible vulnerability disclosure (e.g. alerting the source of the leak in due time) and requiring external review. However, even in a cost/benefit analysis, this argument may fall apart. Some analysis suggests that password research does not, in fact, improve the use of passwords or PSMs, while it does expose the human weakness in the Internet's favoured authentication mechanism.

Other panellists differ, however, and take a more deontological approach to research ethics. Sonia Chiasson argues that all conceivable harm to data subjects must be excluded from the research, and a reasonable effort would be required to acquire their informed consent. The paper does not expand on her idea, though it will be relevant to understand how far the 'reasonableness' standard stretches, and whether it includes opt-out models. Finally, David Dittrich takes a care ethics approach and asks researchers to justify their actions while taking into account the perceptions of data subjects, interpret research ethics rules in the context of the research, and balance the risks and benefits accordingly (Egelman, Bonneau, Chiasson, Dittrich, & Schechter, 2012).

4) Benevolent researcher:

The paper "Passwords and the Evolution of Imperfect Authentication" (Bonneau, Herley, & Stajano, 2014) puts some of the justifications and promises of password research into perspective. The authors argue that there is a lack of evidence that strict password-composition policies (and resulting password strength managers) do, in fact, produce more secure passwords by internet users. They reference Egelman et al.'s paper that show that password-composition policies do not necessarily produce stronger passwords (Egelman, Sotirakopoulos, Muslukhov, Beznosov, & Herley, 2013). Further, they reference Weir et al.'s paper that suggests that the metrics used to measure password strength are overestimated by an order of magnitude (Weir, Aggarwal, Collins, & Stern, 2010). These findings question the value of the justifications by researchers using password dumps, specifically whether their studies do, indeed, produce a great social good in the form of improved password strength managers. This may not be as straightforward as the research community claims.

Description of the paper

The chosen paper for the first case study fits into the tradition described in the previous sections. Wang et al state that the aim of their study is to:

1. “[q]uantify the security risks introduced by password reuse and modification [...]” and
2. “[...] develop a training-based algorithm to guess a user’s password based on their leaked one.” (p. 1)

To achieve these aims, the researchers gathered password dumps from 107 online services, which were leaked between 2008 and 2016. Their initial research database contained 497,789,976 passwords of 428,199,842 unique users. However, for the purpose of this study, the researchers could only use the username and password combinations of persons who included in several of the separate datasets to comprise the research database. This resulted in a final research dataset of 28,836,775 users containing 61,552,446 passwords.

The researchers provide the following table, which outlines their source data:

Category	#Plain PWs (#Datasets)	Top 3 Largest Datasets
Social	286M (7)	Myspace, VK, LinkedIn
Adult	75.2M (9)	Zoosk, Mate1, YouPorn
Game	40.8M (13)	Neopets, 7k7k, Lbsg
Entertain	30.7M (4)	Lastfm, Swingbrasileiro, LATimes
Internet	16.4M (18)	000webhost, Comcast, Yahoo
Email	9.6M (3)	Gmail, Mail.ru, Yandex
Forum	1.1M (25)	CrackingForum, Abusewith.us, Gawker
Shopping	340K (12)	RedBox, 1394store, Myaribags
Others	210K (7)	Data1, Data2, Data3
Business	10K (9)	Movatiathletic, Hrsupporten, 99Fame
Total	460M (107)	Myspace, VK, LinkedIn

Table 1: Categories and statistics of collected datasets.

Figure 4. Password dumps used in case study 1.

The researchers analysed the chosen passwords in a variety of ways. First, the researchers quantified the number of times individual data subjects reused a password across services. Second, the researchers applied eight common password transformation rule combinations to guess a users' password on an online service in their own dataset. These include, for example, adding digits to the tail end of a password phrase (e.g. "Internet2018"), replacing letters with similarly looking digits (e.g. "1nt3rn3t"), or using sequential keys on a keyboard (e.g. "qwerty" or "12345"). Third, the researchers divided the users up in demographics, to analyze whether different professions or countries of origin have distinct usage patterns. Finally, the researchers built an algorithm to efficiently predict passwords based on their analysis of the frequency a transformation rule is applied by people using the online services.

Wang et al. conclude in their paper that the majority of Internet users reuse their passwords, or modify them to some predictable extent, based on one of the eight transformation rules. Of the users in their database, 38% had reused a password and 21% modified their password on at least one other online service. The researchers find that this pattern is consistent across demographics. Therefore, persons who may wish to intrude into online services have a relatively high likelihood of guessing a password based on the most common transformation rules. In fact, in their tests, the researchers' algorithm guessed 30% of passwords within 10 attempts, and 46.5% of passwords within 100 attempts. The researchers thus raise the alarm in their paper, where they state:

“Our result sheds light on the emerging security threats introduced by massive data breaches, and calls for more effective tools to secure users' online accounts and digital assets.” (p. 1)

The authors thereby claim to study a topic of great social importance. People's identities, personal and professional lives are managed to a large extent through online services, which are typically accessed by memorised passwords. If it is, indeed, relatively easy for third parties to access individual accounts on online services by guessing these passwords based on transformation rules, trust and reliance on online services is likely to diminish.

The researchers justify their use of password dumps by the importance of their research topic. In the paper, they state that they worked closely with their IRB to obtain approval for their research. More precisely, the researchers make the following statements in their ethics section:

1. “First, we only analyze datasets that are already publicly available. Analyzing such data does not add additional risks other than what already exist.”

2. “Second, these datasets are also publicly available to potential attackers. Failure to include the data for research may give attackers an advantage over researchers that work on defensive techniques.”
3. “In the past decades, leaked password datasets have been extensively used in academic research [8, 9, 13, 17, 32–34] to develop security mechanisms to protect users in the long run.” (p. 2)

These justifications are analysed through the lens of applicable research ethics frameworks and the conceptual framework in the following sections.

Common Rule analysis

This part of the chapter answers the second research question for this case: “Which concepts and principles of research ethics frameworks are challenged specifically by technical Internet research with unsuspecting data subjects?” In this part, I retrace the decisions the IRB would have made when reviewing this study and I critically examine their reasoning. While this analysis is speculative—the authors did not respond to a request to be interviewed—the decision charts developed by the US Federal Government’s Office for Human Research Protections (OHRP)⁶⁸ to aid navigation through the Common Rule do not leave much room for speculation.

Wang et al. state that they worked closely with their local institutional review board (IRB) to obtain approval for their research on password dumps. The statements and omissions in their section on “ethic guidelines” give a clear indication how the IRB

⁶⁸ See <<https://www.hhs.gov/ohrp/regulations-and-policy/decision-charts/index.html>>.

reasoned to approve this research. Wang et al.'s description of their research ethics considerations is as follows:

Ethic Guidelines. Our work involves analyzing leaked datasets that contain sensitive information. We have worked closely with our local IRB and obtained the approval for our research. Our study is motivated by the following considerations. First, we only analyze datasets that are already publicly available. Analyzing such data does not add additional risks other than what already exist. Second, these datasets are also publicly available to potential attackers. Failure to include the data for research may give attackers an advantage over researchers that work on defensive techniques. In the past decades, leaked password datasets have been extensively used in academic research [8, 9, 13, 17, 32–34] to develop security mechanisms to protect users in the long run. (p. 2)

This statement is illuminating regarding the considerations of the IRB. First, the researchers state that they “obtained the approval for our research”, which indicates that the research project was subject to research ethics review according to §46.102(d) (research) and §46.111 (approval) of the Common Rule. Second, the mention of risk to data subjects is an indication that the IRB determined that the research includes human subjects according to §46.102(f). Third, the approval for the research is an indication that the study was not exempt from research ethics review, for example by classifying password dumps as mere existing and publicly available data, which would trigger an exemption to review based on §46.101(b)(4). Finally, neither the section on research ethics, nor the rest of the paper mentions informed consent. It is therefore plausible to deduce that the IRB exempted the researchers from gaining the informed consent of the millions of data subjects in their study according to the criteria presented in §46.116(d). All four criteria are analysed in the next subsections.

The next part discusses the ambiguity that remains when applying these criteria to the case study paper.

Research and approval

The paper by Wang et al. is inspired by practical problems of securing individual online accounts with a memorised password mechanism. It contributes to a longer stream in computer science scholarship, as described in the previous section. This section demonstrates how this technical research fits within the definition of *research* provided by the Common Rule in §46.102 (d):

“[...] a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge [...]”

Systemic investigation

The stated intentions and design of Wang et al’s paper suggest that their work is indeed a systematic investigation which is designed to contribute to generalizable knowledge. Beyond stating that “there is an urgent need to systematically assess users’ password reuse and modification patterns and quantify the security risks”, the authors describe the three criteria of the systematic approach to their data collection of password dumps:

1. The password dumps had to be available on public forums or websites,
2. They had to contain email addresses to link users across services,
3. The authors excluded databases that contained only salted hashes⁶⁹, due to the difficulty of recovering passwords from this form of security.

⁶⁹ <https://crackstation.net/hashing-security.htm#salt>

Contribution to generalizable knowledge

The researchers state two clear goals in their paper, both of which contribute to generalizable knowledge. First, they aim to “[...] understand how users reuse and modify their passwords across online services at a large-scale.” Second, the researchers “[...] quantify the security risks introduced by password reuse and modifications after data breaches.” The paper present four distinct findings:

1. Password reuse and modification are still very common.
2. Sensitive online services have a high ratio of reused and modified passwords.
3. Users still reuse the already-leaked passwords for years after the data breach.
4. Modified passwords are highly predictable.

Finally, the paper states that the project “[...] was supported by NSF grant CNS-1717028”, which is a grant from the US federal government’s research funding agency. This establishes that the proposed study is, indeed, subject to the US Common Rule and IRB review. This case study does not raise any issues for the definition of *research* as provided by the Common Rule.

Approval

The criteria for IRB approval of research stipulated in §46.111(a)(2) requires the IRB to conduct a harm and benefit analysis. The ethics section of the case study paper justifies the research based on the benefits of it social importance. The researchers state that their research data is equally available to malevolent persons who may intentionally cause harm to the data subjects. If benevolent researchers—such as themselves—would not be allowed to carry out this study and share the results, it would leave the malevolent attackers at an advantage in the ongoing race to secure digital data, access to services, and the trustworthiness of digital identities. Finally,

the researchers appeal to precedent, stating that their study is one in a longer tradition of password research which uses data from illicit sources. To weigh the benefits and harms, it is necessary to have a conceptualisation of both. This is addressed in the third part of this chapter titled “Structured case research cycle analysis”.

Human Subjects

Should data subjects in a password dump be considered as human subjects, as defined by the Common Rule? According to §46.102(f), a research project involves human subjects when its researchers collect data about:

“[...] a **living individual** about whom an investigator (whether professional or student) conducting research obtains

(1) Data through **intervention or interaction** with the individual, or

(2) **Identifiable private information.**”

Living individual

While some of the Internet users, whose email addresses and passwords were leaked, may have deceased in the period between the time of setting up their account and the use of the databases for research purposes, the vast majority are likely to be *living individuals*.⁷⁰

Intervention, interaction, identifiable private information

The researchers did not set up an experiment to collect primary research data through intervention or interaction. However, their datasets contain millions of records that are considered to be identifiable private information. First, email

⁷⁰ My own login credentials to several online services have been exposed in the password dumps used by the researchers. This was checked on the website <https://haveibeenpwned.com/>

addresses are typically individually identifying when they contain the name of a person. The US National Institute of Standards and Technology, for example, lists email addresses as an example of a type of information that is qualifies as *personally identifiable information*.⁷¹ Similarly, the EU considers emails to be personally identifiable when they are in a format such as [name.surname@company.com](#).⁷² The password dumps used by the researchers will likely contain a large amount of identifiable email addresses.⁷³

Second, the Common Rule describes private information in §46.102(f) as:

“Private information includes information [...] has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public [...]”

A password is inherently *private information*, because the Internet user is expected to memorise and keep secret the string of symbols which has the specific purpose of individual and unique authentication for an online service. The Internet user can reasonably expect the service provider to not make the password public, because it would render their authentication procedure exposed and meaningless if their password was not secured and treated as confidential. The data subjects in the password dumps used by Wang et al. are therefore human subjects in the sense of the Common Rule. This case study does not challenge the definition of human subjects.

⁷¹ See <<https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf>>. Other Federal agencies and departments also refer to email addresses as examples of personally identifiable information, see for example the Department of Labor’s Guidance on the Protection of Personal Identifiable Information: <<https://www.dol.gov/general/ppii>>.

⁷² See <https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en>.

⁷³ My personal email address, benzevenbergen@gmail.com, is included in several of the password dumps used by the researchers, for example.

Exemptions from review

Before the study is scrutinised by the full committee, the IRB must determine whether exemptions apply under §46.101(b). Exemptions can be triggered when the proposed study is conducted as part of an educational practice, if the study observes public behaviour, when a study uses only existing data, if it evaluates a public benefit or service programme, or if the study involves a food quality or consumer acceptance study. If the proposed study fulfils any of these criteria, the research is exempt from review and does not require further scrutiny.

The researchers state twice in the research ethics section that their research data is “publicly available.” A previous section of this chapter demonstrates that the public availability of this data is typically one of the justifications for researchers to use password dumps in their research projects. Interestingly, the IRB evidently did not consider (all of) the password dumps used in this research to qualify for the exemption of research ethics review under §46.101(b)(4). Otherwise, the study would have been exempted from research ethics review and there would not have been a need for the sentence “[w]e have worked closely with our local IRB and obtained the approval for our research” to be included in the paper. The relevant exemption reads as follows:

(b) Unless otherwise required by department or agency heads, research activities in which the only involvement of human subjects will be in one or more of the following categories are **exempt** from this policy: [...]

(4) Research, involving the collection or study of **existing data**, documents, records, pathological specimens, or diagnostic specimens, if these sources are **publicly available** or if the information is recorded by the investigator in such a

manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

The two relevant criteria in this provision are that the data must be “existing” and “publicly available.” In a strict interpretation of the letter of the law, password dumps would qualify for both criteria. Password dumps consist of data that were created by users of online services before the breach occurred and are therefore “existing.” While it can be argued whether password dumps are “publicly available”, from a technical point of view it is relatively easy to access these leaked databases. The next part of this chapter analyses the case study through the lens of the conceptual framework and assesses the extent to which scholars have previously addressed the definitions of ‘existing data’ and ‘publicly available’ with regards to data available on the Internet.

Waiver for informed consent

Obtaining the informed consent and processing responses would have been a major part of this study, had this not been waived by the IRB. This section retraces the steps the IRB would have taken to arrive at this conclusion. The section highlights two moments at which the IRB may have made a misinformed decision: (1) determining that the data subjects' rights would not be adversely affected if informed consent was not sought and (2) determining that the proposed study would not be feasible without a waiver for informed consent.

In principle, the Common Rule prescribes that researchers should attain informed consent from their data subjects in §46.116. This requirement can be waived by an

IRB according to §46.116(d) if three conditions are met, and providing the IRB with an option on an optional fourth condition:

§46.116(d): An IRB may approve a consent procedure which does not include, or which alters, some or all of the elements of informed consent set forth in this section, or waive the requirements to obtain informed consent provided the IRB finds that:

- 1) The research involves no more than minimal risk to the subjects;
- 2) The waiver or alteration will not adversely affect the rights and welfare of the subjects;
- 3) The research could not practicably be carried out without the waiver or alteration; and
- 4) Whenever appropriate, the subjects will be provided with additional pertinent information after participation.

We can deduce from Wang et al.'s paper that they did not seek informed consent from data subjects within the password dumps, so the IRB must have determined that the proposed study fulfils at least one of the first three criteria listed above. Presumably, the IRB did not require the researchers to provide additional pertinent information to data subject, either. This section describes why this decision to grant a waiver for informed consent may not be in compliance with the letter of the law.

Wang et al. amassed as many password dumps as possible. They then made a selection based on specific criteria for this study. There was no interaction between the data subjects and the researchers. Neither did the data subjects know—nor did they consent to—that their data was or would be used for this kind of research.

The first condition requires the study to pose no more than minimal risk to data subjects. With regards to password dumps, the researchers argue that "[a]nalyzing such data does not add additional risks other than what already exist." While this may seem justifiable at face value, the next section analyses how the considerations of Internet researchers in the social sciences have considered the requirement of "no more than minimal harm" to be problematic in their work, and how that relates to this case study paper. It is not relevant for this section.

The second condition asks whether a waiver would "adversely affect the rights and welfare of the subjects". It could be argued that, since the study is unlikely to pose additional risks to subjects per se, it is also unlikely to affect their welfare. However, by processing the personal and identifiable data of research subject for another purpose—that is to say, research rather than identification—the rights of data subjects can be adversely affected. For example, the European data protection laws (either the directive 95/46/EC or later the regulation 2016/679) provide a set of legal grounds for processing (such as informed consent or compliance with a contract). Similar provisions will be relevant for data subjects from other jurisdictions. Since the password dumps include data from people around the world, the processing of their data (unique identifier and password) are violations of the rights as data subjects. It can thus be argued that waiving informed consent and processing the data in password dumps does, in fact, adversely affect the rights of at least some of the data subjects.

The third condition requires that the research would not be practically feasible without a waiver for informed consent. The IRB may have erroneously determined that this condition is fulfilled. It may, indeed, be burdensome to request the

informed consent of the 28.8 million unique Internet users included in the primary research data set. However, the researchers state that the password dumps do include the email addresses of users as their unique identifiers across services, so direct communication with data subjects would—in theory—be possible. Sending 28.8 million emails, recording the responses, and filtering out the passwords of data subjects who do not agree to participate in the research, or who did not respond, is indeed a monumental administrative task. But it is not impossible, and it would not necessarily halt the research. The third part of this chapter addresses an alternative option that was not explored by the researchers or the IRB.

Structured case research cycle analysis

This part of the chapter answers the third research question for this case: “How might these concepts be reinterpreted to be useful and appropriate for the ethical assessment of technical Internet research?” It is therefore one of the core parts of this thesis. This chapter is the first case chapter of the ‘structured case’ research cycle, and second iteration on the conceptual framework developed in Chapter 3.

Continuing from the previous chapter, the emerging interpretations of concepts and research themes are summarised in four tables at the end of the chapter. The next case study chapter builds further on the results of this part, which contribute to the final conceptual framework in chapter 6.

This section is somewhat repetitive by necessity. It analyses the extent to which the information gathered in the chapter is relevant for its structure, and integrates the interpretation of concepts by social scientists and computer scientists with regards to using password dumps through the lens of the Common Rule framework.

The starting point of this part is the ambiguity identified in this chapter so far. Not every concept of the conceptual framework is included here, because the case did not identify all as being problematic or ambiguous. The previous section identifies the following ambiguity:

- **Privacy:** The nature of password dumps as digital data that is existing and publicly available data,
- **Autonomy:** Justifications to waive informed consent because of its impracticality,
- **Minimal risk of harm:** The concept of harm, specifically "minimal harm", when using a database comprised of several other databases that have been made available through illicit means.

Each subsection is a deeper assessment of the concepts of the research ethics framework. This section does not explicitly present a new conceptual framework. This task is left to chapter 6, which is a deeper analysis of the emerging interpretations of this thesis. For now, each of the ambiguous and new concepts is analysed and taken into account in the next case analysis.

Are password dumps “existing” and “publicly available” data?

As described in the introductory sections of this chapter, passwords are the dominant authentication mechanism for Internet services. With usernames and passwords of another person, an attacker can imitate another’s online identity and access sensitive details of that persons' life. Users are typically required to memorise their username and passwords and are encouraged not to share them. The same is true for Internet services, of which Internet users expect a high level of security for

their passwords. Understood as such, a password is a particularly sensitive piece of information, which is part of a key mechanism to make Internet services function adequately for individuals.

A password dump is typically comprised of millions of username and password combinations, exposing them to anyone who downloads the files. While some services do send notifications of breaches to their users, not all users will necessarily change their password after it has been exposed. Further, as the paper by Wang et al. demonstrates, many Internet users reuse passwords or have a predictable system for their passwords across Internet services to memorise them more easily. This may enable hackers or interested parties to learn the system of a user to more easily guess their password for another service that has not been breached.

Wang et al.'s study would have been exempt from research ethics review, had the IRB determined that password dumps are *publicly available*. A previous section notes that research papers which use password dumps typically claim that they only use data that is already publicly available, thereby alluding to the exemption in the Common Rule. However, given the IRB's review and permission to conduct the research, they cannot have accepted this reasoning for an exemption. This section analyses the public nature of password dumps through the lens of the conceptual framework chapter, which confirms the IRB's assessment.

Password dumps are available for download on several websites or forums on the Internet. A simple search engine query is enough to find and download password dumps that contain millions of entries.⁷⁴ Given the secrecy required to make

⁷⁴ For example, an article titled "Today I Am Releasing Ten Million Passwords" is one of the first search results using common search engines: <https://xato.net/today-i-am-releasing-ten-million-passwords-b6278bbe7495>

passwords a useful authentication mechanism, and the ability to expose intimate details of the bearer's life, the mass violations of privacy posed by these dumps is apparent. It is an open question, though, to what extent such data qualifies for the research ethics exemption of *publicly available* data in §46.101(b)(4) of the Common Rule. A further complication arises because the researchers do not interact with the data subjects, or intervene in their life. For example, Christina Allen argued that the Internet has blurred the distinctions between what is public, and what is private information (Allen, 1996). Erin Kenneally and Dave Dittrich have found that no clear guidance nor consensus exists for computer science researchers making use of publicly available data from an illicit source (Kenneally, 2015b; Dittrich et al., 2012).

It is, however, possible to articulate a reasonable norm against the reuse of password dumps. Previous experiences of Internet researchers, their assessment of their responsibilities, and the reasoning of their critics can be applied to password dumps as follows:

The level of accessibility of information on an Internet service prescribes the perception of privacy of Internet users, and their motivation to share information (King, 1996; Pittenger, 2003). Internet users share usernames and passwords with Internet services with the understanding that both parties will keep the information secret to serve the higher-order goal of maintaining the individual authentication mechanism, and to limit access to private data and an online identity. The expectations of privacy by Internet users dictate the role a researcher can play within a particular online space, and, by extension, with regards to data that originates from such a space (Buchanan, & Ess, 2008). The technical ability and capacity to collect and analyse this data does not confer upon the researcher the right to reuse this data.

Instead, Jim Thomas hints at a context-sensitive and responsive approach, claiming that their technical knowledge "confers upon researchers an obligation to exercise special caution when information is obtained from informants who do not know they are the subjects of a study [...]." Following this reasoning, a responsive attitude would dictate that the violations of Internet users' expectations and perceptions of privacy with regards to their password data indicate that the data should therefore not be reused for research.

Waiver for informed consent requirement

Requiring researchers to gain the informed consent for participation of their data subjects is commonly accepted as a mechanism to safeguard participants' autonomy. It allows data subjects to choose whether to participate, understand how to withdraw from a study, and protects vulnerable groups with diminished autonomy, for example through a proxy person who is authorised to give consent on another's behalf. As described in the second chapter of this thesis, however, some scholars question to what extent informed consent can safeguard the autonomy of vulnerable groups, who are considered to have a diminished autonomy in the first place.

The second chapter of this thesis describes a range of issues with the mechanism of informed consent for Internet research that have been noted by scholars. First, the data subject is typically not clearly identified, if at all. The distance between researcher and data subject—both in terms of geography and perception by the researcher—makes the human data subject and their local context somewhat invisible to the researcher. Frankel and Siang ask: "[...] how valid is such consent when the age, competency, or comprehension of the potential subject is unknown?" (Frankel, M. S. & Siang, 1999, p. 10). If their identity remains unknown, it

will be impossible to make a judgement of their local risks of harm in participating in a study. Further, it will not be possible to fulfil the obligations of communicating the ways in which a data subject can withdraw from a study. For these reasons, Nathaniel Poor and Roei Davidson decided against using data from a leak in a different context, because they concluded that it was unacceptable to use peoples' hacked and leaked data without their consent (Poor & Davidson, 2016).

Justification to waive informed consent requirement

The question is, however, to what extent the IRB was justified in waiving the requirement to gain informed consent from unsuspecting Internet users, whose usernames and passwords were implicated in a password dump that were used for the research study? Each condition to waive the requirement to gain informed consent is assessed separately:

*(1) The research involves **no more than minimal risk** to the subjects;*

The third chapter of this thesis describes how social scientists and technical Internet researchers have grappled with the 'risk of harm' concept in their disciplines. In summary, social scientists describe the concept of harm as being "ambiguous" when applied to online experiments. Internet users who may be geographically dispersed can be implicated in research in a variety of ways, making the probability and magnitude calculation complex to the point that it may not be possible to generalise the "daily life" risk standard to the population within a dataset. The levels of harm to data subjects may differ significantly between the data subjects contained within a dataset, so a single standard has not been considered to be appropriate. Indeed, to fulfil the "respect for persons" principle on the Belmont Report, scholars argue that researchers must take into account the local context of data subjects to make a fair

assessment of the risks they may face. This requirement also features in the ethics of care literature, as described in the second chapter.

The study by Wang et al. did not target Internet users directly, though their usernames and passwords are implicated in technical Internet research. It will be a difficult task to determine with accuracy whether the reuse of password dumps has exposed Internet users to indirect or further harms that are outside the control of the researchers. Persons implicated in a password dump may not experience direct physical or financial harm per se, though resulting effects can be considered as moral harms or violations of users' legal and moral rights. It may further be argued that by publishing their paper, which details the different password dumps in existence, the researchers benefit personally from an illicit act and may have led other parties to find these databases of which they were previously unaware. These considerations are discussed in the next sub-section.

A responsive approach through the lens of the ethics of care is useful. Following the research themes developed in the second chapter, researchers would need to articulate a responsive approach to the needs of their data subjects with whom they have entered into a relationship by using their data. This requires researchers to understand the needs and expectations of data usage by their data subjects, as well as the established rules, etiquette and guidelines of the venue or site from which the data is collected. Because it may not be feasible to assess the context of each of the millions of records, researchers would need to generalise the perceptions and expectations of their data subjects, who have provided the source data for the password dumps. A strong argument exists to conclude that unsuspecting Internet users would consider their privacy to be violated if their passwords were reused.

(2) *The waiver or alteration will not adversely affect the **rights** and **welfare** of the subjects;*

Rights: As previously noted in this chapter, the combination of emails as usernames and passwords are most likely protected *personal data* or *personally identifiable data* according to data protection laws of the various jurisdictions around the world.

These laws typically require the processors of such data—which are the researchers using password dumps in this case—to gain the informed consent from their data subjects as one of the grounds which makes the processing lawful (among others, which are not relevant for this thesis). By not gaining the informed consent from the individual data subjects, researchers may well be violating their data subjects' rights. Not only is a violation of rights a violation of §46.116(d)(2) of the Common Rule, which states that “[t]he waiver or alteration will not adversely affect the rights and welfare of the subjects”, it may also be a *moral harm*. The different conceptualisations of *harm* are discussed in a further sub-section.

Welfare: The welfare of vulnerable groups is most likely to be affected when in oppressive or hierarchical relationships. Internet users whose authentication credentials have been implicated in password dumps can be categorised as a vulnerable group for several reasons. First, their sensitive authentication data have been leaked from Internet services and have been made available on forums for other people to use beyond their initial expectations of how this information would be kept secure and secret. Unless the password is changed manually once it has been compromised, this potentially exposes their data and communications to others and may allow them to control their online identities. Including the source of password dumps in research papers may alert potential attackers to their existence and unintentionally inspire further malicious actions. Second, the data subjects have not

given their informed consent, likely have no knowledge of the research being conducted with their data, and have no ability to protest or withdraw from the study. Finally, it is unclear whether Internet users belong to other, existing vulnerable groups, or how the concept of vulnerableness is to be understood in the local context of the various individuals implicated in the password dumps. In summary, researchers will typically not be able to adequately assess the extent to which their research affects the welfare of its data subjects, or to what extent special precautions are needed to ensure their welfare is not adversely affected.

(3) The research could not practicably be carried out without the waiver or alteration

The third chapter describes that the technical internet research community often requests waivers for informed consent, because it is not possible or practical to carry out their research without it. Typically, the identity of Internet users whose data is implicated in a research data set will be untraceable and thus their identity remains unknown. This makes it largely impossible to communicate with these data subjects. Further, it is considered to be a difficult task to explain in layman's terms how the technical research is conducted. Keneally and Dittrich note that the requirement for a waiver of informed consent is stricter than it merely being difficult or inconvenient.

Similar issues have been faced in bioethics, where Luke Gelinas et al. argue that the “impracticality” condition should be understood as (i) imposing prohibitively high costs of time or economic resources, or (ii) threatening to undermine the validity of the proposed research. In addition, though, they state that research without consent is only justifiable if (a) it does not infringe or data subjects’ rights, or (b) an infringement of the rights of data subjects does not have substantial moral gravity and the violation is outweighed by the social value of the research (Gelinas,

Wertheimer, & Miller, 2016). Gelinias et al. suggest that applying this framework requires ethics reasoning when they state that:

"The hard ethical work will be to determine when research without consent violates participant rights, when any such rights have substantial moral weight, and when it is fair to say that it is impracticable to obtain consent." (p. 43)

For this case, it is relevant to assess whether gaining the informed consent of data subjects (and thereby being responsive to the need of autonomy) would (1) pose a prohibitively high cost of time and resources, and whether (2) the violation of rights in data protection laws do not carry substantial moral gravity and therefore do not weigh up against the social value of the research project.

(1) Time and resources

Wang et al. state several times in their paper that the identifiers in the research dataset are email addresses. While it would be a monumental administrative task to provide relevant information, request the consent, record the responses, and act on decisions whether to participate or not, it does not seem impossible to communicate with their data subjects in principle.⁷⁵ The question for this case study then becomes whether the size and scale of a research database with existing data relating to identifiable human subjects can be considered *too* impractical to require a researcher to gain the informed consent before using their data for research, and it thus being practically infeasible to carry out the proposed research.

While it could be argued that the administrative burden would be too great to record millions of informed consent notices, it does not seem like the researchers explored a

⁷⁵ The technical challenges to sending millions of emails have been addressed by several organisations. See for example <<https://drivy.engineering/sending-mass-emails/>> or services such as <<https://www.sendblaster.com/>>.

relatively simple compromise: sending data subjects an email with an explanation of the study and offering the option to opt-out of the study. Such a strategy could be automated by offering participants the opportunity to enter their email address in a form on a dedicated website, for example, which could remove their passwords from the research data set prior to analysis. Opt-out strategies are often used in medical science (see for example (Vellinga, Cormican, Hanahoe, Bennett, & Murphy, 2011; Junghans, Feder, Hemingway, Timmis, & Jones, 2005)) and do not necessarily lead to a large drop out rate (Hewison & Haines, 2006). Approached in this way, it does not seem that gaining informed consent through an opt-out procedure would be too burdensome or impractical for researchers.

(2) Moral gravity of violation and social value of research

Weighing the moral gravity of a data protection violation against the social value of potentially strengthening password mechanisms is a complex task that is prone to justified disagreements. A lot more information would be needed to construct an argument in favour or against the moral significance of these values. The aim of this thesis is, however, to illuminate the difficulty of applying existing research ethics frameworks to technical Internet research, and to identify the weaknesses. For this case study chapter, it is sufficient to state the tension between the values, and the complexity of making a thorough justification either way. The next subsection does inform this consideration to some extent.

Conceptualising harm

The discussions in the previous section about a minimal risk harm with regards to Internet users—who are implicated in password dumps that are subsequently reused for research purposes—do not offer satisfying conclusions. To assess whether

these Internet users are harmed, it is necessary to dive deeper into the complexities of the concept of harm. This section presents a brief—not conclusive nor exhaustive—overview of ‘harm’ as it has come to be understood in philosophy, and showing why Ben Bradley concludes “that the best course of action is to avoid appealing to the notion of harm at all in our moral theorizing” (Bradley, 2012, p. 410-411).

Risk of harm is an important concept in philosophy, in part to distinguish wrong actions from right ones. In professional medical ethics, the *Hippocratic Oath* contains the part-sentence “[...] into whatsoever houses I enter, I will enter to help the sick, and I will abstain from all intentional wrong-doing and harm, [...]” (Jones, 2013, p. 616). In the development of utilitarianism, John Stuart Mill’s harm principle published in 1859 in his influential work *On Liberty* states “[t]he only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others” (Mill, 1966). Further, the precautionary principle—an interpretation of the German environmental policy principle *Vorsorgeprinzip* —is an expression of a need to anticipate harm before it occurs. A popular interpretation is contained in the *Wingspread Statement on the Precautionary Principle* in 1998, formulated as “[w]here an activity raises threats of harm to the environment or human health, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically” (Ashford et al., 1998).⁷⁶ The European Commission has declared the precautionary principle to be relevant for decision-making with regards to the “[...] risk of adverse effects to the environment, human, animal or plant health”, especially to “[...] build a common

⁷⁶ Online publication. Available at: <https://www.gdrc.org/u-gov/precaution-3.html>.

understanding of how to assess, appraise, manage and communicate risks that science is not yet able to evaluate fully [...]”.⁷⁷

It is clear that the concept of risk of harm has been—and still is—a foundational philosophical concept in the development of professions, policy, and science. But what actually is harm? Ben Bradley states that “[d]espite the importance harm is supposed to have, almost nobody bothers to say what it is” (Bradley, 2012, p. 391). It could refer to physical pain, financial loss, reputational damage, a barrier to new opportunities, violations of rights, or violations of other normative values, such as autonomy, dignity, privacy, or even the process of dehumanisation. To interpret the concept of harm in this case, it is useful to assess the concept of harm in the context of reusing password dumps for research purposes through the lens of taxonomies of harm created by philosophers, and how they are applied in reasoning methodologies such as consequentialism and deontology.

Donald Warwick proposed a three-tiered taxonomy, specifically for social research (Warwick, 1982):

- 1) Harms to research participants,
- 2) Harms to society,
- 3) Harms to researchers and the research profession.

a) Harms to research participants

Regarding the harm to research participants, Alasdair MacIntyre proposes that these can be divided into three further categories:

- A harm to a person's interest,

⁷⁷ Online publication. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52000DC0001&from=LT>.

- A wrong to a person,
- A moral harm to a person.

The concepts of *interest*, (*moral*) *wrong*, and *moral harm* do not have clear and distinct definitions, though they are used often used in established philosophy literature.

According to MacIntyre, a harm is rectifiable or commensurable, while a wrong, by its nature, cannot be compensated for (MacIntyre, 1982⁷⁸, see also Holden, 1979). A harm to a person's interest forms the basis for tort law, while a wrong—such as lying—is typically accepted as something to be tolerated in human society. A harm to a person's interest is typically a moral wrong, too, though it is narrower in scope. Eriksson and Helgesson describe moral harms—in the context of biobank research—when they state that researchers “may treat an individual in an illegitimate way even if they are not exposed to any risk of physical, psychological, social or economic harm” (Eriksson & Helgesson, 2005, p. 1072). They mention violations to autonomy, privacy, personal integrity, and dignity, which are closely related to the requirements of “respect for persons” of the Belmont report. Further, they describe moral harms to groups, stating:

“[...] groups can also be wronged, even if no single individual is pointed out specifically. For instance, groups are treated badly when they are used for research because of their weak position or inability to protest.” (p. 1073)

In the ethics of care literature, a moral harm is described as the “severing of valued connections” (Kittay, 2011, p. 53). The various descriptions of moral harms are closely related to the inclusion of vulnerable groups in research. Vulnerable persons are conceptualised as those clearly lacking decisional capacity or those deprived of

⁷⁸ Page number uncertain due to Covid-19 related library closures.

their autonomy. Research ethics literature considers the inclusion of vulnerable groups—defined as such—to be exploitative and a violation of the autonomy and dignity, and thus a violation of moral harms. Internet users, whose usernames and passwords have been exposed in password dumps and whose data is then subsequently included in research projects, can therefore be considered as a vulnerable group who lack decisional capacity that are devoid of autonomy with regards to this sensitive data. Including their data in a research project can therefore be considered to be a moral harm to the Internet users.

b) Harms to society

In philosophy, the concept of harm is typically considered with regards to the individual. Lord Patrick Devlin introduced the concept of ‘injury to society’ into legal theory, arguing that the law as a whole gains its legitimacy with its function to protect society from harm. He states the following:

“[...] But the true principle is that the law exists for the protection of society. It does not discharge its function by protecting the individual from injury, annoyance, corruption, and exploitation; the law must protect also the institutions and the community of ideas, political and moral, without which people cannot live together” (Devlin, 1959, p. 125).

Harms to society can be manifold, and it is not a simple task to speculate what the harms society experiences by researchers reusing data found in password dumps. Wang et al. argue—along with other researchers who reuse password dumps—that their research benefits society by developing more robust password strength managers. But could it not also be argued that their methods of analysis described in the research paper and the resulting algorithms developed by the researchers set a precedent and inspire other malevolent persons who may have not previously

considered infiltrating Internet service accounts, or lacked the technical know-how to carry out their unlawful actions with others' usernames and passwords? It is difficult—if not impossible—to quantify and weigh to what extent the research based on password dumps is, indeed, beneficial or harmful for the trust society places in the password mechanism to keep their information secure.

c) Harms to researchers and the research profession

A more obvious harm to society is the harm to the research profession, its resulting loss of public trust, and new restrictions on academic freedom. Kenneth Howe and Michele Moses, in discussing Warwick's taxonomy of harm, argue that "[I]ncreased public mistrust of scholarly work negatively affects the public perception of researchers and their activities, as well as the possibilities for funding and support of important research endeavors" (Howe & Moses, 1999, p. 27). If research professionals are permitted to commit moral harms to unsuspecting Internet users, the profession may be perceived as somewhat vulturous, scavenging on the fruits of crime and even benefitting from it personally and professionally. For example, if a researcher is promoted or recognised for a study that has been made possible by harming data subjects, it sets a precedent and creates a research culture that makes this behaviour acceptable and will likely incentivise other researchers to also harm other unsuspecting data subjects for their personal and professional gain. It is with good reason that research institutions prohibit further uses of the outcomes and data of studies that has been condemned as "unethical", or those that are in violation of research ethics principles. Henry K. Beecher argues that the duty to prohibit use and publishing of research that violates ethical principles weighs stronger than the

potential benefits of such research. He states, with regards to sadistic human experimentation by Nazi doctors:⁷⁹

“Even though suppression of such data (by not publishing it) would constitute a loss to medicine, in a specific localized sense, this loss, it seems, would be less important than the far reaching moral loss to medicine if the data thus obtained were to be published” (Beecher, 1966, p. 77).

The outcome of this reasoning is that harm must be understood in a variety of ways, beyond simply assuming that it is self-evident what harm is in a given context.

While it may be enticing to do away with the concept of harm in moral theorising, as suggested by Ben Bradly, the Common Rule prescribes a harm assessment. It may therefore be more favourable to consider the concept of *minimal risk*—defined as “the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life [...]” in §46.102 of the Common Rule—more holistically, to include moral harms and rights violations.

Conceptual Framework: 2nd Iteration

Human subjects	
Relational approach to morality	Data subjects: • Technological mediation ◦ Relationship established through the use of networked technology, ◦ Enrolment into relationship may not be voluntary, ■ Consent technically not necessary for interaction or intervention. • Distance principle ◦ Disintermediated relationship through affordances of networked technology, ◦ Increased experiential distance of researcher ◦ Researchers and data subjects not easily identifiable ◦ Data subject is <i>distant other</i>. • <u>Passive data collection.</u> ◦ <u>Received from third party.</u> ◦ <u>No intervention or interaction.</u>

⁷⁹ I note that the harm done to victims of Nazi doctors is incomparable to unsuspecting Internet users whose passwords have been exposed, though the comparison on the use of data from illicit sources is relevant.

	Society: • Hierarchical societal position of researcher ◦ Academics enjoy a powerful position, ◦ Technologists shape human communications and interaction, • Random people can be implicated in an Internet research study by virtue of using the Internet. • Power relationship to data subject, ◦ Potentially oppressive relationship, ◦ Data subject is vulnerable to the decisions made by researchers. • <u>Data subject's data is technically publicly available.</u> ◦ <u>Not publicly available in the meaning of the concept.</u> Research profession: N/A
Contextually-oriented deliberation	Research context: • Assessing the implications of a disintermediated relationship is based on a lack of knowledge. Data subject context: • Difficulty or impossibility to know local context of data subjects, • Replace dichotomous human subject concept with continuum based on a variable such as identifiability or potential risk of harm.
Responsiveness	Identify: • Human ownership or use of implicated devices, Understand: • Data collection through intervention or interaction with devices may implicate humans, Interpret: • Sensitivity to the offline context within which the data subject is likely nested.
Emotions as moral tools	Positive: • Beneficial research, ◦ By benign researchers. • Possible to understand and meet the expectations of Internet users. Negative: • Humans may be used as a means to an end, ◦ If local context is not identified and understood, ◦ And researcher is not responsive to their specific needs. • Research can be considered as a form of surveillance

Minimal risk of harm	
Relational approach to morality	Data subjects: • <u>Harm can materialise in a variety of ways</u> ◦ <u>Interests</u> ◦ <u>Wrongs</u> ◦ <u>Moral harms</u> Society: • <u>Weigh harms against the social importance of the research</u> • <u>Injury to society</u> ◦ <u>Committing harms sets precedent for:</u> ■ <u>Inappropriate acts by researchers.</u> ■ <u>Inappropriate acts by malevolent parties.</u> • <u>Effect of harms to relationship with society.</u> ◦ <u>Knowledge gain to research disciplines, vs.</u> ◦ <u>Higher-order loss of trust to research discipline.</u> Research profession: • <u>Loss of public trust in research discipline may affect availability of funding.</u>
Contextually-oriented deliberation	Research context: • A single 'daily life' standard does not exist, ◦ Variables in the typical calculation of the magnitude of risk multiplied by its probability are too complex and vague to be meaningful, • Researcher may not be able to assess whether their actions constitute a minimal risk of harm. • <u>Vulnerability:</u> ◦ <u>Data subjects may be vulnerable persons according to existing definitions.</u> ◦ <u>Data subjects may be vulnerable as a result of being implicated in the research.</u> Data subject context: • Harms are highly contextual, • Data subjects can be situated in wide-ranging contexts with undefined sociopolitical parameters, • Research may create new risks of harm in the data subject's context: ◦ By disregarding local (informational) sensitivities, ◦ By being insensitive or ignorant to existing power relations, ◦ By violating local laws. • <u>Violations of moral harms include:</u> ◦ <u>autonomy</u> ◦ <u>privacy</u> ◦ <u>personal integrity</u> ◦ <u>dignity</u>
Responsiveness	Identify: • Human ownership or use of implicated devices, • Context-specific harms, • Unsuspecting data subject expectation and perception should be leading: ◦ Researchers to seek out rules, etiquette, guidelines that govern the virtual space or community. • <u>Self-serving arguments:</u> ◦ <u>Will the perceived benefits indeed materialise, or are they hypothetical?</u> ◦ <u>Research does not necessarily deliver the benefits suggested by researchers</u> • <u>Moral harms stemming from vulnerability and ability to object.</u> ◦ <u>Illegitimate treatment even if data subjects are not exposed to any risk,</u> ◦ <u>Groups are treated illegitimately when they are used for research because of their weak position or inability to protest.</u> Understand: • Researcher may not be able to understand or generalise the data subjects' needs, interests, and the context-specific risks of harm, • Harms may only become apparent after the data collection or publication of the research, • <u>Interests of colleagues may be harmed if (moral) harms occur.</u> • <u>Publication of research may incentivise others researchers to use data from illicit sources, creating potential for further harms.</u> Interpret: • Which standard is the researcher imposing on all data subjects? • Interpret minimal risk standard on a continuum, instead of the static and universal 'harm [...] no greater than those ordinarily encountered in daily life' standard, ◦ Differentiate between:

	■ Direct harm, ■ Indirect harm, ■ Potential harm. • <u>Interpret harm beyond monetary loss and physical injury.</u>
Emotions as moral tools	Positive: • Benevolent intentions, ○ <u>Higher-order goal of improving security.</u> Negative: • Possibility of conducting deception research without protocol, • Exploitation of vulnerable persons, ○ who are not aware of the research, ○ who cannot exercise their rights, ○ who are unable to object or protest • <u>Selfishness: Scavenging on the fruits of crime</u> ○ <u>For personal and professional benefit.</u> ○ <u>Risking reputational loss to discipline.</u>

Autonomy	
Relational approach to morality	Data subjects: • Unknown or misleading identities, ◦ Researchers, ◦ Data subjects. • Status of relationship is unclear if identities are unknown, • <u>Possibility to communicate establishes potential relationship.</u> ◦ <u>Researchers should explore possible channels of communication with data subjects.</u> ◦ <u>Presence of email addresses in dataset allows communication.</u> Society: • Researchers can enrol Internet users seemingly at random ◦ Diminishing the ability of all Internet users to choose to take part in research, • Lacking clarity whether informed consent establishes autonomy or a relationship meaningfully in research at all, • People's interactions with those who have consented, can be implicated, • Ability by researchers to alter psychological state of Internet users • Ability by researchers to change Internet users' environments. Research profession: N/A
Contextually-oriented deliberation	Research context: • Can researchers infer the intentions and expectations of users to provide existing data? ◦ <u>Higher-order reason why users provided this data dictates its context and reason for researcher to process further</u> • The researcher may be required to make an effort to make themselves known to data subjects, • Difficulty or impossibility to identify vulnerable persons or groups within the research population, • Communication about foreseeable risks is not feasible without a contextual analysis, • Notification of rights (e.g. to withdraw) or debrief after deception may not be possible. • <u>If informed consent is not feasible.</u> ◦ <u>Consider alternatives (e.g. opt-out).</u> ◦ <u>Weigh moral gravity of potential rights violation against the social value of research outcomes.</u> • <u>Other applicable laws may prescribe necessity for informed consent.</u> Data subject context: • Difficult or impossible to understand data subject context, • <u>Researcher has responsibility to assess the intention of data subject:</u> ◦ <u>E.g. to keep data secret.</u> • <u>Responsibility to assess whether the reuse of data exposes data subjects to new risks (indirect) harms.</u>
Responsiveness	Identify: • <u>Starting point: gain informed consent</u> ◦ Are data subjects in a position to give their informed consent? ◦ Uncertainty whether research subjects understand what they consent to, ◦ Competence to provide informed consent may vary, ◦ Vulnerability status (decisional capacity) of research subjects. ◦ <u>Legal requirements for informed consent.</u> ◦ <u>Can the data processing and purpose be explained to data subjects in layman's terms?</u> ◦ <u>Assess burden of consent</u> ◦ <u>Consider alternatives, e.g. proxy person, opt-out</u> • <u>Alternative: waive informed consent</u> ◦ <u>Overly complex if data subjects are dispersed around the world.</u> ◦ <u>Consider:</u> ■ <u>High costs of gathering informed consent (time & resources).</u> ■ <u>Moral gravity of the potential violation of data subjects' rights.</u> ■ <u>Threat to undermine validity of research.</u> Understand: • <u>Impossibility to make a judgement of needs if the identity of data subjects is unknown.</u> ◦ <u>Obligation to understand the expectations of data usage by data subjects.</u> • <u>Assess potential harms beyond legal infringements, monetary loss, and physical harm.</u> • <u>Assess vulnerability of groups in research population.</u> • <u>Can the data subjects protest or withdraw from the data collection?</u>

	Interpret: • <u>Reasonable effort to obtain informed consent.</u> • <u>Explore information provision in layman's terms.</u> • <u>Can informed consent alone safeguard the autonomy of vulnerable groups?</u> • <u>Consent via proxy: is the person giving consent (e.g. from website or platform owner on behalf of their users) legitimised to do so?</u>
Emotions as moral tools	Positive: • Respecting autonomy through engagement with data subjects, Negative: • Outside interaction with a person's device may be a violation of rights, ◦ <u>Violation of rights can (also) be considered as a moral harm.</u> ◦ Potential violation of the freedom from being subjected to research, • <u>Violation of autonomy to use data without consent.</u> ◦ Violations of dignity, • <u>Recklessness:</u> ◦ <u>Potential to create further harms beyond the control of researchers.</u> • <u>Exploitation:</u> ◦ <u>No knowledge of research by data subject.</u> ◦ <u>Data subjects are used as a means for research because of their weak position or inability to protest.</u>

Privacy	
Relational approach to morality	Data subjects: - Relationship to researcher established through networked technology connectivity, - The identity of Internet users is intertwined with their devices due to increasing sociotechnical convergence, - Re-identification in reused datasets may establish a relationship to previously anonymous data subjects, <u>Data points may be inherently identifiable</u> <u>Data points may enable communication (e.g. email addresses).</u> - Those with technical capability to collect data, establish the relationship. Society: - Researchers can in theory collect research data about any Internet user, - Without the Internet users need for knowledge or consent for data collection. <u>Data sources may not be intended for public use.</u> <u>Even though they are technically accessible, freely or for a fee.</u> <u>Passive data collection: purpose for sharing data should be leading.</u> - The power to collect data establishes a power hierarchy, - A form of surveillance. <u>Violations of privacy can be justified by higher-order purposes.</u> <u>E.g. improving security mechanisms for data on the Internet.</u> <u>Justifications must be robust.</u> <u>Negligible risk to data subjects.</u> Research profession: <u>Data from illicit sources are extensively used in technical Internet research.</u>
Contextually-oriented deliberation	Research context: - Distinction between public and private domains are blurred. <u>Obligation to reason assess data availability beyond mere legal compliance (e.g. data protection, information privacy, research ethics).</u> <u>Intended accessibility of data indicates perception of privacy by data subjects.</u> <u>Higher-order goals and purpose of data usage to infer expectations and perceptions.</u> Data subject context: - For passive data collection, the reuse of data in research context may not be perceived as appropriate. <u>Consider intention by data subject to keep data secret</u> <u>Data unlikely intended for research purposes or public exposure</u> <u>Data can come back to haunt data subjects.</u>
Responsiveness	Identify: - User expectations and perceptions of privacy in their online context, <u>Respect high-order goals of privacy as part of technology functions.</u> - For passive data collection, identify the purpose for which data is originally shared, - Potential for re-identification after anonymization of dataset. Understand: - Sensitivity of data for data subjects, - Obligation to exercise special caution, - When data subjects do not know they are participating or data is collected about them, <u>No authoritative guidance exists for data use from illicit sources.</u> - Lines between data collection, illicit intrusion into devices, and deception, - For passive data collection from an illicit source, the responsibility for researchers may be unclear. Interpret: - Special caution when collecting data on unsuspecting data subjects, <u>Technical ability and capacity to collect data does not confer the right to use this data.</u> <u>Context-sensitive and responsive approach requires special caution.</u> - Tension: Data utility vs. purpose limitation, <u>Do advances for science justify speculative ends of research?</u> - Determine level of privacy to be respected based on: - Group accessibility, - Perceived privacy of data subjects.

	○ <u>User expectations and perceptions of privacy dictate the role a researcher may play in an online space.</u> ● Offline analogies should not replace assessment of user expectations.
Emotions as moral tools	Positive: ● Control of data, ○ To maintain dignity. Negative: ● Potential for surveillance, ● Loss of privacy, ○ Reported feeling of being dissected like laboratory animals, ○ Internet users become a means to an end for researchers. ● Feeling of being violated.

(5) Case Study 2: Censorship measurement

Introduction

The second case study chapter analyses a research project titled *Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests*, authored by Sam Burnett and Nick Feamster at Georgia Tech's School of Computer Science. Feamster moved to Princeton University while the study was in progress. The paper is suitable as a second case study for this thesis, because it contains an active data collection methodology by the researchers. They devised a technical mechanism by which information flows are initiated on the device of unsuspecting data subjects. This is in contrast to the previous case study, which used only existing data for its study. Burnett and Feamster's paper was originally submitted to the Internet Measurement Conference (ACM IMC) in May 2014, but received a sceptical response from the programme committee based on research ethics concerns. This chapter focuses on a second iteration of the paper, submitted to the ACM Sigcomm conference in 2015.

This study came to my attention in a variety of ways. First, I was introduced to the authors of the paper while it was under review at the ACM IMC 2014 conference, because the programme committee's deliberations on research ethics were unclear and inconclusive at the time. Second, the programme committee at the ACM Sigcomm conference in 2015 asked me to organise a workshop at their conference, to explore the ethical challenges presented by this type of research. This workshop was

the final workshop organised as part of the Ethics in Networked Systems Research project.⁸⁰

This chapter, like the previous case study analysis, is divided in four parts. The first section provides a description of the case study paper and situates it within its wider research domain. The second section analyses the research paper through the lens of the applicable research ethics framework, the US Common Rule that was applicable at the time the research team would have gone through the application procedure (ie. before the revision of 2018 came into effect).⁸¹ The aim of this section is to identify the areas where the research ethics framework leaves ambiguity or may be inappropriate. The third section assesses the extent to which the ambiguities that surface in the Common Rule analysis can be addressed by the existing scholarly debates covered in the conceptual framework, or to what extent new issues arise. This section iterates on the developing conceptual framework, which informs the final analysis of the next chapter. Finally, the fourth part iterates on the tables, which summarise the further development of the conceptual framework.

Contextualizing the case study paper

The paper by Burnett and Feamster is exemplary for the ongoing experimentation with innovative data collection methods that enable a thorough measurement of the Internet infrastructure and the behaviour of the various components on the network. This first section begins by situating the case study paper with its research domain. First, the academic debate around computer science research that measures the

⁸⁰ See <<https://www.networkedsystemsethics.net>>.

⁸¹ The Common Rule has been amended and revised, though it has not yet become fully effective due to ongoing delays. See <<https://www.federalregister.gov/documents/2018/06/19/2018-13187/federal-policy-for-the-protection-of-human-subjects-six-month-delay-of-the-general-compliance-date>>.

networks and objects that comprise the Internet, is discussed. The second subsection reviews technical Internet censorship measurement research. Finally, the third subsection describes the case study paper and highlights how statements made in the paper connect to the existing literature.

Internet measurement

History and objective

The Internet has been characterised as a "network of networks" (Salus & Vinton, 1995; Ballot, Montreuil, & Meller, 2014) and a "bundle of complementary components that are used in different proportions, in different contexts, no single metric is sufficient" (Lehr, 2012, p. 8). The architecture of the Internet was built taking into account values such as simplicity, flexibility, decentralisation, and interconnectedness, which allowed for vast experimentation and innovation, as discussed in more detail in Chapter 3. As a result, the Internet developed as a complex digital computer network that is difficult to map, study, and fully comprehend. In fact, the Internet has not been mapped holistically (see for example Bu & Towsley, 2002; Eriksson, Dasarathy, Barford, & Nowak, 2012; Zhang, Liu, Massey, & Zhang, 2005), so we do not actually know what the topology of the Internet looks like at any given moment in time. This lack of knowledge has become a problem increasingly recognised by policymakers, researchers, and engineers tasked with administering the Internet, who can never fully know how the components of the network interact precisely. The immense scale of the overall network leads to highly variable behaviour of the various components in their separate technological, social, and political contexts, which makes it difficult to characterise the network accurately (Rabinovich, M. & Allman, 2016).

To address this problem, researchers began to measure and keep track of the ever-growing network already in the early stages of Internet development, in order to better understand the behaviour of decentralised components. Out of these efforts, the academic discipline called 'Internet measurement' was established, which is now a sub-discipline of 'computer science' and 'network engineering'. Research in this field aims to measure, for example, the topology of the Internet, how the various decentralised components of the Internet behave, and the performance of newly introduced protocols. Uncovering this knowledge about the Internet and collecting real-time data is arguably essential for the functioning and sustainability of the network as a whole (Lehr, 2012). Scholars involved in Internet Measurement make use of the affordances of the Internet's architecture to observe the behaviour of the networked devices on a large scale without the active participation or knowing cooperation of users. Such studies can be staggeringly large in scale and make use of a variety of data sources, both collected with active and passive methodologies.

Technologies and methodologies

The dynamic nature of the Internet means that research methodologies and data collection for Internet Research must keep up with the rapidly changing environment. Already in the year 2000, Internet Measurement pioneer Kimberly Claffy noted that Internet traffic, technology, and network topology change faster than Internet measurement studies can adapt (Claffy, 2000). Rabinovich and Allman note that research methods in Internet measurement became increasingly complex and innovative along with the further evolution of the Internet as a complex system (Rabinovich, M. & Allman, 2016).

Researchers now typically collect data from a wide variety of vantage points on the Internet network to gain accurate measurements. Experimental mixed-methods approaches improve the inferences about the behaviour of components on the Internet, and are able to attribute actions to actors without their knowledge by initiating data flows from another person's devices. The next few paragraphs describe some common methodologies and their differences.

Data collection for Internet Measurement can be active or passive. *Active* measurement involves initiating an information flow or an action in a pre-defined way between nodes connected to the Internet. Measurements using this technique are designed to infer behavioural or topological properties of networks, the reach of a given stream of traffic, and other observed network-related reactions. A commonly used methodology for active measurement is network scanning for security assessment of networks using the ZMap research tool, which allows performing fast scans on an Internet-wide scale (Dumeric, et al. 2013). Other methods of Internet measurement may exploit known software or security 'bugs', that allow researchers to take control of a device, albeit mostly to a limited extent. Finally another methodology called 'idle port scanning' uses 'side-channel attacks' to quite literally 'bounce' initiated information flows off an Internet-connected device to make measurements of how it connects or relates to other devices (Ensafi, et al. 2010). These side-channel measurements are different from exploitation of security bugs, because the researcher doesn't take control over devices in any way. The bugs that allow side channels are common to many different implementations and may be inherent to some protocol specifications.

Passive measurement involves watching—or surveilling—existing Internet traffic and collecting the available data. The variables measured can be the same as those measured actively—such as speed, latency, path—but are not initiated by the researcher. This type of measurement can be used to determine, for example, how quickly a user can reach online services, or whether specific websites respond to requests, and which variables may have contributed to observed behaviours.

Another popular methodology to collect data about the Internet is called a "cross-origin requests", or "cross-origin resource sharing." In short, web browsers allow code in websites to embed objects from other servers to be requested as the website is loaded and displayed to the user. This method has been exploited by several researchers, for various ends. For example, Karir et. al use an embedded script which triggers cross-origin requests to measure the reachability of IPv6 addresses⁸² using a large number of clients (Karir, Huston, Michaelson, & Bailey, 2013). Gummadi et. al were able to measure latency between DNS name servers⁸³ by embedding cross-origin requests to websites based on other DNS servers and recording the time it takes to receive a response (Gummadi, Saroiu, & Gribble, 2002). Lam et. al use cross-origin requests to show how they may be exploited as an information security vulnerability, whereby an Internet user's browser can be co-opted to be included in malicious activity, such as "denial of service attacks, worm propagation and reconnaissance scans" (Lam, Antonatos, Akritidis, & Anagnostakis, 2006, p. 221).

⁸² The Internet Protocol (IP) is a key communication protocol that allows the routing of information flows to and from specific locations on the network. IPv6 replaced IPv4 with more characters in the addresses, when it became clear that the amount of IPv4 addresses would be exhausted due to the high number of components and devices being connected via the Internet. See the announcement of the IPv6 standard here: <https://www.internetsociety.org/blog/2017/07/rfc-8200-ipv6-has-been-standardized/>.

⁸³ The Domain Name System (DNS) name servers are applications that provide translation services between numerical IP addresses and alphabetical URLs.

Concerns

The benefits—or even necessity—of Internet measurement are clear: to maintain a sustainable technological growth of the Internet, the current state and the effect of technical and policy interventions must be monitored, evaluated, and understood. Scholars have noted some serious ethical challenges to the technologies and methodologies used for the automated data collection that support Internet measurement, however. The uptake of the Internet in many aspects of human life have caused human values to be affected by Internet experimentation and where collecting data on the infrastructure is no longer an isolated activity (Claffy, 2000). When a computer network—such as the Internet—is the subject of research, its connected users may thereby also be considered as research subjects (Wright, 2006; Rabinovich, M. & Allman, 2016). In fact, Internet Measurement researchers may purposefully use the devices and Internet connections of unsuspecting Internet users as the nodes and vantage points in their data collection processes, without their knowledge that they are participating in a research study.

Rabinovich and Allman point out that passive measurements of Internet traffic may be considered as surveillance and a violation of privacy. Further, collecting this data may well be a violation of data protection laws (Rabinovich, M. & Allman, 2016). They note that active measurements will alter the behaviour of an Internet user's device and potentially change the context in which they operate. This has the potential to change an Internet user's informational, political, or even physical environment, to the point of inadvertently incriminating users in activity where they did not intend to play a part. After all, the researcher may be unaware what active experimentation on another person's device will affect in practice, or what the knock-on effects are. In Rabinovich and Allman's view, it is imperative for

researchers to actively gain some foresight into the side-effects of their measurements on the lives of their implicated data subjects and mitigate foreseeable harms in their research design. This is further complicated by cultural and language barriers, which need to be surmounted to understand the concerns of activating nodes and vantage points in various locations in the world.

The third chapter of this thesis already described Partridge and Allman's suggestion for a framework for the risk of harm assessment. They differentiate between direct harm, indirect harm, potential harm, and the nature of the harm caused by Internet research. *Direct harm* can be inflicted on people or their systems by actions designed to target them as a data subject specifically by the researcher. *Indirect harms* occur when the research targets technical systems for research, but humans are implicated in an unforeseen manner. *Potential harm* can occur when the actions of the researchers set up the possibility of harm only when other events outside of their control occur (Partridge & Allman, 2016).

Censorship measurement

Censorship practices and studies thereof have a longer history than the existence of the Internet. Measuring censorship has, to some extent, also been a topic of study in a variety of academic disciplines, such as political science, history, literature, art history, cultural studies, and media studies. In these areas, scholars typically study the aims, motivations, scope, types of content, and the legitimacy of censorship policy.

Internet censorship measurement is a part of the academic community that focuses specifically on technologies that limit or control access to information on the Internet. Typically, these limitations are imposed by State actors, though they may also be the result of private and corporate decisions (which, in turn, may or may not be due to State pressure). Internet censorship measurement studies attempt to gain knowledge about how a variety of actors censor information on the Internet. Elahi and Goldberg propose a commonly referred to definition of Internet censorship when they describe it as follows:

"Internet censorship is the intentional suppression of information originating, flowing or stored on systems connected to the Internet where that information is relevant for decision making to some entity" (Elahi & Goldberg, 2012, p. 3)

The knowledge gained from Internet measurement studies may be used by others (or the researchers themselves) to develop circumvention tools to subvert the restricted flow of information. Further, the knowledge created may feed into studies of the other relevant academic disciplines.

Internet censorship techniques

From a technical perspective, Internet censorship can be instituted on three levels. First, *client-based censorship* is software that restricts access to information on the device that is connected to the Internet. For example, parental controls at home or companies and institutions that employ a firewall to limit the access to certain websites. Second, *server-based censorship* occurs when a censor impairs access to content by removing it or otherwise limiting access to it. This type of censorship is commonly the result of complying with a formal request or State order, though may also be the result of a policy or decision by the owner of the server. Finally, and most relevant for this chapter, *network-based censorship* occurs when the communication

mechanics between a network client and the server are disrupted. This is the most efficient method of censorship, as it gives the censor wider coverage, controlling a high number of communications (Aceto & Pescapé, 2015).

A wide range of network-based censorship techniques have been uncovered by researchers. Below is a brief overview of some of the most well-known methods, along with a description of their functionality:

- **BGP tampering:** The Border Gateway Protocol (BGP) is the algorithm that coordinates packet-forwarding (the basic functionality of how data is transferred across the Internet) between various administrative boundaries on the Internet. Tampering with the BGP in a specific region can result in traffic being routed through a specific designated server (which may act as a so-called *choke point*), or the Internet being unreachable for users in that territory (for example the events during the Arab Spring⁸⁴).
- **DNS tampering:** The Domain Name System (DNS) translates the symbolic name (in the form of a Unique Resource Locator — URL) of a server hosting a resource to their specific Internet Protocol (IP) address. If a DNS server is tampered with, the censor may refuse to allow the translation of URLs to specific IP addresses, and instead return forged error messages to the user.
- **TCP connection disruption:** The Transmission Control Protocol (TCP) is the algorithm that establishes a connection between two endpoints on the Internet (ie. the client and server). A censor may alter the algorithm in a specific region

⁸⁴ See the article "How Egypt shut down the internet":
<<https://www.telegraph.co.uk/news/worldnews/africaandindianocean/egypt/8288163/How-Egypt-shut-down-the-internet.html>>.

to effectively disallow connections to specific IP addresses on which the targeted content is hosted.

- **TLS tampering:** The Transport Layer Security (TLS) is a protocol that provides cryptographic end-to-end communications security over networks, commonly used on the Internet. A censor may alter the protocol by forcing it to approve a false security certificate, leaving the user under a false pretence of security.
- **HTTP tampering/proxy filtering:** The HyperText Transfer Protocol is the underlying protocol used by the World Wide Web that defines how information are formatted and transmitted. If the TLS connection has been tampered with, a censor may hijack the communication and send other information to the end users, such as the well-known "404 page not found" error.
- **Packet filtering:** A censor can place a surveillance and censorship device at specific, strategic point on the Internet network to covertly drop data packets when they are in transit. For example, it may drop packets that are related to a particular IP-address, thereby sending a 'connection timed out' error to the end user requesting content from that server or URL. Further, a censor may drop packets when they contain specific keywords. Packet filtering may also be used to throttle internet traffic, lowering the quality of service for certain servers to the point of inaccessibility.⁸⁵ This type of censorship is harder to measure than other, more obvious types of connection failure.

⁸⁵ See, for example: <<https://www.sciencedirect.com/topics/computer-science/stateful-packet-filtering>>.

Methodologies and technologies

The Internet protocols and standards have been developed to enable a decentralised and transparent flow of information. John Gilmore stated in 1993 that "[t]he Net interprets censorship as damage and routes around it."⁸⁶ Today, however, many governments actively try to censor information on the Internet using the methods outlined above. These attempts to restrict the flow of information violate the core principles of the Internet architecture, disrupt the functioning of network protocols, and mangle the various components that constitute the Internet in some way.

Jones et. al state that the Internet censorship measurement sub-discipline is "driven by the desire to have reliable, continuous, and comprehensive empirical data about the nature of censorship around the world" (Jones, Ensafi, Feamster, Paxson, & Weaver, 2015, p. 17). There are two main elements researchers in this field use to accomplish this objective:

First, researchers deploy vantage points on networks to which they do not themselves have direct access. This can be achieved by sending researchers to foreign locations or asking peers in other areas of the world to deploy the measurement software on their behalf. Well known examples of this software are provided by the Open Observatory of Network Interference (OONI) project,⁸⁷ the open source Centinel software,⁸⁸ or by the PlanetLab project.⁸⁹ Alternatively, researchers may send active probes using existing devices on the Internet network to request to load sources that may be censored, and return the measurement results to

⁸⁶ See <<http://www.chemie.fu-berlin.de/outerspace/internet-article.html>>.

⁸⁷ <<https://ooni.torproject.org/>>.

⁸⁸ <<https://pypi.org/project/centinel/>>.

⁸⁹ <<https://www.planet-lab.org/>>.

the researchers. Examples of this are the Spooky Scan,⁹⁰ the use of side channels (Ensafi, Knockel, & Alexander..., 2014), and the paper for this case study chapter.

Second, the researchers require a list of websites that are known to be—or likely to be—censored. The measurement software is then programmed to target those specific websites to collect data. These URL lists can be obtained from organisations such as the Open Network Initiative,⁹¹ the Herdict project,⁹² the GreatFire project based in China,⁹³ and previously the Filbaan project in Iran.⁹⁴

Concerns

Internet censorship measurement studies exist within a tension between gaining valuable information about the censorship activities of governments, while potentially incriminating or putting the persons conducting the measurements at risk of retaliation from those governments, whether they were part of the measurements or not. The case study paper is an exemplary study that raises all the common concerns of Internet measurement research. These will therefore be discussed in detail in the next subsection.

Description of paper

The paper titled *Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests* exemplifies the need for researchers conducting Internet censorship measurements to innovate their experimental data collection mechanisms. This

⁹⁰ <<https://www.cs.unm.edu/~jeffk/spookyscan/>>.

⁹¹ <<https://opennet.net/>>.

⁹² <<http://www.herdict.org/>>.

⁹³ <<https://en.greatfire.org/>>.

⁹⁴ <<http://web.archive.org/web/20141219030923/https://www.filbaan.net/>>.

chapter focuses on the version of this paper published as part of the ACM Sigcomm 2015 conference in London, because in computer science the programme committee of a conference acts as the peer-review mechanism, as well as—to some extent—the institutional research ethics review board.

Aim

The authors describe the aim of their study in various parts of the paper. The first aim is to demonstrate "the feasibility of a global-scale deployment with a pilot study and an analysis of potentially censored Web content" (p. 653, 2015 version). Their system, 'Encore', is a technical data collection system that continuously collects measurements from a global set of vantage points. Encore can thereby provide an up-to-date view of Internet censorship worldwide, illuminating what content is censored, in which locations, and for whom.

The second stated aim is to provide data and inferences about Internet censorship and "report them to a central authority for analysis." In the authors' view, Encore can give a broader picture of global Internet censorship than other, existing tools. When used in combination with other, existing tools for censorship measurement, Encore fills an important gap that enables researchers to draw inferences about Internet censorship techniques employed by governments with a higher confidence than previously possible.

The authors describe some further advantages of Encore over other Internet censorship tools: (1) It does not require the recruitment of individual volunteers to deploy the software and is thus easily scalable. (2) The authors claim that a censor

would not be able to detect and disable the Internet measurements conducted, since they come from many devices and not just from one, central source.

Technology

This section describes the flow of data initiated by Encore. On a high level, the technical functionality of Encore is described by the authors as follows:

"We present Encore, a system that harnesses cross-origin requests to measure Web filtering from a diverse set of vantage points without requiring users to install custom software, enabling longitudinal measurements from many vantage points."
(p. 1, 2014 version)

The authors provide the figure below.

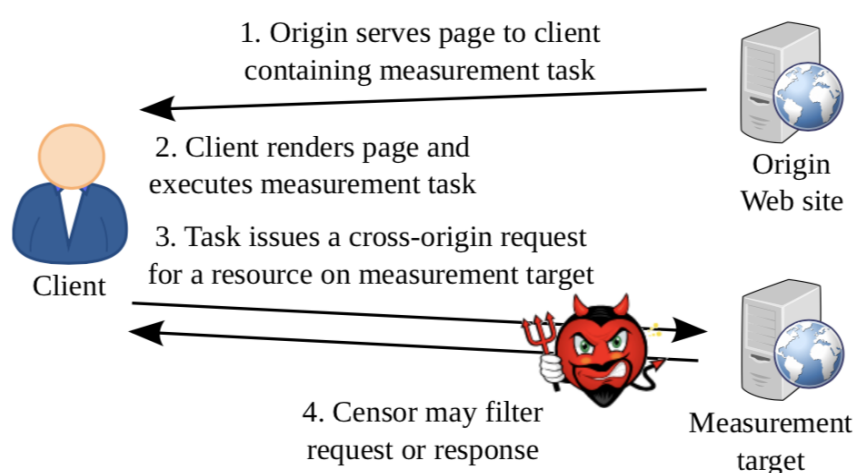


Figure 5. Encore induces browsers to collect Web filtering measurements by bundling measurement tasks inside pages served by an origin Web site.

For clarity, the flow of data is described with two prior steps (A and B), in addition to the four outlined in figure X.

(A) The webmaster of a website ("Website X") installs this code snippet provided by the Encore researchers.

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

This code directs the user's device to a widget on the website that initiates Internet censorship measurement requests to a selection of URLs when Website X is rendered (this is the measurement task). The widget is invisible to the user of Website X, since its width and height are set to "0" pixels, and the display of the widget is disabled.

(B) An unsuspecting Internet user types in the URL, or clicks on a link to the URL, which redirects to Website X.

(1) "Origin serves page to client containing measurement task". The full code of Website X is transferred to the user's device, including the invisible widget.

(2) "Client renders page and executes measurement task". As the received code renders in the user's browser, their device connects to Encore's server (on URL <https://encore.noise.gatech.edu/task.html>)

(3) "Task issues a cross-origin request for a resource on measurement target". Encore's server initiates a series of cross-origin requests from the user's device to

URL's that are contained on a predefined URL list located on Encore's servers. The user's device acts as a vantage point for Encore's Internet censorship measurement.

(4) "Censor may filter request or response". If the code of websites designated on the URL list, to which measurements were sent, fail to load, the authors assume the content is censored in some way. The measurement results are sent from the user's device back to Encore's server, along with the unsuspecting Internet user's IP address. A limitation of Encore is that it can "[...] detect *whether* certain URLs are filtered, but it cannot determine *how* they are filtered" (Burnett & Feamster, 2015, p. 654).

The data flow described above is the innovative element Encore contributes to the Internet censorship measurement literature. The authors state that "[t]he cornerstone of Encore's design is to use information leaked by cross-origin embedding to determine whether a client can successfully load objects from another origin." In addition, the authors state some requirements beyond these steps. They do not intend to "[...] compromise security of the page running the task". Further, the byte size of their requests must be small enough to "not significantly affect perceived performance, appearance, or network usage."

URL lists

In the first version of Encore's measurements, the researchers used "[...] a list of several hundred 'high value' URLs curated by Herdict and its partners." In the second version of the paper, which is the focus of this chapter, the authors added a sentence, stating that "[c]urating accurate and appropriate lists of potentially censored URLs is beyond the scope of this paper and an active research area."

Further, in the second version, the authors importantly state that "[b]ecause measuring Web filtering may place some users at risk, we only measured Facebook, YouTube, and Twitter." The authors have thus reduced the number of URLs to be tested from several hundred to only three in the second version of the paper.

Recruitment

Recruitment for Internet censorship measurement studies typically requires researchers to liaise with individuals who have access to a specific region's Internet network. These persons are then instructed to install and deploy the custom made software. The authors describe the creation of vantage points for Encore as follows:

"In contrast, Encore recruits a relatively small number of webmasters and piggybacks on their sites' existing Web traffic, instantly enlisting nearly all of these sites' visitors as measurement collection agents." (p. 660, 2015 version)

Encore makes use of the users of a small number of websites that have chosen to integrate Encore's code snippet into their HTML code. If installed on a sufficiently diverse number of websites, a large and dispersed group of unsuspecting Internet users will inevitably execute the Encore code, which gives the researchers a potentially worldwide and real-time view of Internet censorship.

The authors argue that webmasters are incentivised to install this code, if they want to contribute to Internet censorship measurement and transparency on government web-filtering practices. Further, the authors suggest that, in future, they could offer a reciprocal agreement, whereby the participating webmasters' URLs are added to Encore's URL list. In return for giving access to their users' devices as vantage points, the webmaster may discover where their websites may potentially be censored.

Data collection

The second version of the Encore paper states that 17 webmasters had participated in the initial data collection from May 2014 until January 2015. In this period, the researchers had triggered Internet censorship measurements from 88,260 distinct IP addresses, which were spread across 170 countries. The authors report the most measurements originating from the following countries: China, India, the United Kingdom, Brazil, Egypt, South Korea, Iran, Pakistan, Turkey, and Saudi Arabia. All of these countries are known to practice some form of Internet censorship.

Research ethics concerns

This section presents the research ethics concerns raised by the authors. Interestingly, the concerns in the second version of the paper had evolved substantially from the concerns raised first version, so both versions are presented here for reference. The research ethics decision-making by the authors' university IRB and the conference programme committee are discussed in the next part of this chapter.

The first version of the Encore paper includes a 'Security Considerations' section. This section describes the security measures taken to avoid external interference with measurement results, as well as a short sub-section titled 'User privacy and protection.' The authors make four observations:

- 1) The content to which Encore sends requests through unsuspecting Internet users' devices is potentially sensitive and may be incriminating,

- 2) Even if it would not be explicitly unlawful to try to access this content in the various contexts where Encore will initiative measurements, it may be harmful for unsuspecting Internet users,
- 3) However, the authors argue, that the unsuspecting Internet users have plausible deniability (e.g. in a court of law), because they were not aware of the Internet censorship measurement that were occurring as a side effect of loading another website.
- 4) Still, the authors "recognize that developing and deploying a tool like Encore entails risks that [they] need to understand better."

The second version of the Encore paper contains a more extensive and in-depth discussion of the research ethics concerns. The authors note that the Encore study "raises ethical questions that researchers in computer science rarely face and that conventional ethical standards do not address." They make an argument for why they chose not to request the informed consent of unsuspecting Internet users, whose devices they used as vantage points for their study. In typical consequentialist reasoning that is common among engineers, the authors present the emerging ethical issues as a difficult balance between benefits and potential harms (Lloyd & Busby, 2003).

The benefits are stated as follows:

- An innovative, efficient, global, and scalable scientific measurement tool,
- New, valuable scientific data for scholarship,
- Access to data collection that was previously prohibitively expensive,
- Ability to measure Internet censorship in broad strokes.

The risks are described as "far more nebulous":

- Potential risks to a complex web of stakeholders,
- Potentially violating laws, though there is a scattered legal landscape regarding the access to censored content,
- Inducing clients to request URLs that might be incriminating or harmful in some countries and circumstances,
- Recognition of the measurement sequence by the censor, who may then block the requests (though it is considered difficult and unlikely).

The authors present the following justifications for their decision to collect the research data without the informed consent of unsuspecting Internet users:

- Informed consent is not required by the Common Rule when it is "prohibitive or impractical and there is little appreciable risk of harm to the subject." They thereby attempt to argue that informed consent is impractical for Internet measurement studies at scale,
- Gaining informed consent "would require apprising a user about nuanced technical concepts" across various languages,
- Further, without justification, the authors argue that "[s]uch burdens would dramatically reduce the scale and scope of measurements" and result in only the existing activists currently engaged with Internet censorship measurement "put[ting] themselves into harm's way to study censorship",
- Gaining informed consent therefore would not, in fact, reduce the risk to users. In contrast, the authors argue, it may even "increase risk to users by removing any traces of plausible deniability".

Common Rule Analysis

This part of the chapter answers the second research question for this case: "Which concepts and principles of research ethics frameworks are challenged specifically by technical Internet research with unsuspecting data subjects?" The authors provide more information about their ethics review process in the second version of the paper. First, they attempted to have their project be reviewed by Georgia Tech's IRB. Then, after a move to Princeton University by one of the authors, they opened a dialogue with Princeton's IRB. Both IRBs declined to review the Encore study, because—according to them—it does not meet the definition of 'human subject research.'

The first version of the paper received criticism on research ethics grounds from the programme committee of the ACM Internet Measurement Conference in 2014 and is rejected on ethics grounds. The programme committee at the ACM SigComm 2015 conference accepted the second version of the paper, though they attached an unprecedented statement to the top of the paper.

This part of the case study chapter describes the reasoning by the IRBs, as well as the reasoning the IRB would have had, if they had taken into account the reasoning by the ACM Sigcomm programme committee. The programme committee raised some further points that are relevant for ethical consideration beyond the letter of the Common Rule, which are mentioned in this part and further described in the third part of this chapter.

IRB review

This part mirrors the second part of the previous chapter. It takes the steps an IRB would chronologically take when reviewing a proposed research study through the lens of the Common Rule.

Research

The paper by Burnett and Feamster is inspired by the problems faced by Internet censorship measurement researchers of not having a baseline dataset of measurements against which new, targeted measurements can be held. The authors describe that Encore was designed as a 'systematic investigation' to 'develop or contribute to generalizable knowledge' in accordance with §46.102(d) of the Common Rule, when they describe their study as follows:

"We present Encore, a system that harnesses cross-origin requests to measure Web filtering from a diverse set of vantage points without requiring users to install custom software, enabling longitudinal measurements from many vantage points", and

"We explain how Encore induces Web clients to perform cross-origin requests that measure Web filtering, design a distributed platform for scheduling and collecting these measurements, show the feasibility of a global-scale deployment with a pilot study and an analysis of potentially censored Web content [...]" (p. 1, 2014 version)

Human Subjects

Reiterating the previous chapters, §46.102(f) of the Common Rule states that a research project involves human subjects when its researchers collect data about:

"[...] a living individual about whom an investigator (whether professional or student) conducting research obtains

(1) Data through intervention or interaction with the individual, or

(2) Identifiable private information."

However, the IRBs at Georgia Tech and Princeton both regarded the study to not involve human subjects as data subjects, and declined to review the study. The authors present the following email exchange with Georgia Tech's IRB representative on their website:⁹⁵

```
Mail from the Georgia Tech Office of Research Integrity
Assurance:

From: "Clark, Melanie J" <Melanie.Clark@gtrc.gatech.edu>
Subject: RE: IRB/Ethics questions
Date: September 17, 2014 at 9:00:18 AM EDT
To: "Feamster, Nicholas G" <feamster@cc.gatech.edu>

Hi Nick,

Thanks for talking about the project. As we discussed this
does not require IRB
review since it doesn't meet the definition of a human
subject. Please delete
the reference to the IRB in the Encore information page.

Thanks again

Melanie J. Clark, CIP
Associate Director
Office of Research Integrity Assurance
Georgia Institute of Technology
404/894-6942
http://researchintegrity.gatech.edu

Our original request:

From: Feamster, Nicholas G
Sent: Monday, September 15, 2014 11:35 AM
To: Clark, Melanie J
```

⁹⁵ See <<http://encore.noise.gatech.edu/irb-mail.txt>>.

Cc: Burnett, Samuel
Subject: IRB/Ethics questions

Hi Melanie,

I would like to talk to you for 30 minutes about this project:
<http://encore.noise.gatech.edu/>

This is not a human-subjects experiment per se, but certain questions have come up, and we'd like to get an authoritative answer from you/IRB about them.

Can I pay you a visit for 30 minutes this week?

Thanks,
-Nick

Similarly, the authors state in the paper that Princeton's IRB also did not consider the study to meet the definition of human subjects. Further, the authors state that the Princeton IRB raised an additional point regarding the potential risk. The authors write:

"It is generally accepted that users already have little control over or knowledge of much of the traffic that their Web browsers and devices generate (a point raised by Princeton's office of research integrity and assurance), which already gives users reasonable cover. By analogy, the prevalence of malware and third-party trackers itself lends credibility to the argument that a user cannot reasonably control the traffic that their devices send. The more widespread measurements like Encore become, the less risky they are for users." (p. 664-665, 2015 version)

Princeton's IRB appears to be giving additional cover to the authors, assuring that—even if the data subjects would be considered *human subjects*—they would not experience more risk of harm than the average Internet user. Indeed, web browsers and digital devices generate data flows that the user is typically not aware of, so Encore's cross-origin requests would not put the user in a position where they may

experience additional harm. Following this reasoning, this would allow the study to be reviewed in an expedited review procedure based on §46.110 and would not require the researchers to obtain the informed consent from data subjects based on §46.116(d)(1) of the Common Rule. So, even if the data subjects would meet the Common Rule's definition of human subjects, the study would only be reviewed expeditiously and no obligations to inform the unwitting participants would be required.

Both refusals to review the study are procedural, leaning on definitions and the letter of the law, rather than the actual impact of the study on the human who own the implicated (or "co-opted") devices. The refusal to review the project through the lens of the Common Rule would make this case study chapter less interesting, were it not for the ACM SigComm 2015 programme committee, had given the Encore study more thought. They allowed the paper to be published in the conference's proceedings, but attached a stern warning at the top of the paper and published a separate review, which addressed the research ethics issues raised and the committee's reasoning process.

Program committee

Studies in computer science are rarely rejected by their institution's IRB. In fact, similarly to the Encore study, technical Internet research is historically considered to be merely technical, and not to include human subjects (Narayanan & Zevenbergen, 2015; Metcalf, J. & Crawford, 2016). The most prestigious venue to publish computer science papers is in the proceedings of conferences. Each iteration of a computer science conferences installs a programme committee, which is tasked with reviewing the submissions. These committees typically do not include scholars with knowledge

of ethics in general or research ethics specifically, though submitted papers are expected to be reviewed and evaluated on ethical grounds. While the merits of research ethics review by conference programme committee in general can be debated, the responses from programme committees to the Encore paper were interesting and unique.

Burnett and Feamster had first submitted the Encore study to the ACM Internet Measurement Conference in 2014. However, after a sceptical response from the programme committee, the authors decided to seek their IRBs evaluation. While the IRB opted not to review the study, the paper was also ultimately not accepted to this conference. Instead, the conference hosted a heated one-hour discussion on research ethics.

The scope of the paper was later reduced to only measuring the censored status of specific URLs, which is presented in the second version of the Encore paper—the case study paper for this chapter. The authors submitted this version to the ACM SIGCOMM conference in 2015. After another internal discussion by the next programme committee, the paper was accepted. However, in an unprecedented move, the programme committee required the paper to be published in the proceedings with the following ‘signing statement’ included at the very top of the paper:

“Statement from the SIGCOMM 2015 Program Committee: The SIGCOMM 2015 PC appreciated the technical contributions made in this paper, but found the paper controversial because some of the experiments the authors conducted raise ethical concerns. The controversy arose in large part because the networking research community does not yet have widely accepted guidelines or rules for the ethics of experiments that measure online censorship. In accordance with the published submission guidelines for SIGCOMM 2015, had the authors not engaged with their Institutional Review

Boards (IRBs) or had their IRBs determined that their research was unethical, the PC would have rejected the paper without review. But the authors did engage with their IRBs, which did not flag the research as unethical. The PC hopes that discussion of the ethical concerns these experiments raise will advance the development of ethical guidelines in this area. It is the PC's view that future guidelines should include as a core principle that researchers should not engage in experiments that subject users to an appreciable risk of substantial harm absent informed consent. The PC endorses neither the use of the experimental techniques this paper describes nor the experiments the authors conducted."

Interestingly, the programme committee notes that the "networking research community does not yet have widely accepted guidelines of rules for the ethics of experiments [...]", which is an indication that they do not consider themselves bound by the Common Rule. The programme committee further raises the concern that the Encore study subjects Internet users to an "appreciable risk of substantial harm", which ideally would have necessitated the authors to obtain informed consent from their participants. One of the programme committee members, John W. Byers, published a public review of the Encore paper, in which he explains the ethics reasoning by the committee in more detail. In addition to the appreciable risk—which Byers describes as potential "severe harm"—he describes how "unsuspecting users around the world must be enlisted to download oft-censored URLs without their informed consent". He further notes the power of cross-origin requests when used in this way, which may set a precedent and thus enable malicious activity, such as enabling new attack vectors for less benign causes (Byers, 2015).

Alternative IRB process

The relevant IRBs and the programme committee at the ACM SIGCOMM conference reviewed the Encore study through different research ethics lenses. The IRBs applied *procedural ethics* with a strict interpretation of the Common Rule. The programme

committee applied an *ethics in practice* approach, assessing the wider effects of the study on affected stakeholders. This resulted in two different outcomes.

Interestingly, the programme committee did not consider itself or the researchers to be bound by the Common Rule. Therefore, their assessment is likely based on a superior technical understanding compared to the IRBs.

First, the IRB did not consider that data collection from devices, which are owned by persons, meet the formal criteria for human subject research. The programme committee did not stop their scrutiny at this definitional hurdle, and considered the impact of the experimental data flows on the device owners. Second, the IRBs considered no additional risks to arise from the Encore data collection mechanisms than those that already exist for Internet users, who are largely unaware of the data flows initiated by their devices in the first place. The programme committee differed again, noting an "appreciable risk of substantial harm absent informed consent" as a result of the Encore data collection mechanism.

This section presents a mock review of the Encore study, taking the stated assumptions of the programme committee and feeding them through a Common Rule review procedure. The section concludes that, given the assumptions of the programme committee, the Encore study would need to be reviewed fully and that informed consent would be required from data subjects. This additional analysis also informs the structured case analysis of the conceptual framework in the next part of this chapter.

Research

First, the programme committee would need to decide whether the Encore study is considered as 'research' as defined in article 46.102(d) of the Common rule. The previous section described this reasoning, whereby an IRB would also have classified as Encore being designed as a "systematic investigation" to "develop or contribute to generalizable knowledge." The programme committee would have come to the same conclusion.

Human subjects

To reiterate, §46.102(f) determines that a research project involves human subjects, when its researchers collect data about:

“[...] a living individual about whom an investigator (whether professional or student) conducting research obtains

(1) Data through intervention or interaction with the individual, or

(2) Identifiable private information.”

This section addresses these criteria in turn.

Living individual

Encore “piggybacks” on Internet users accessing websites that had the experimental widget installed, invisibly to the user. It can be argued whether the researchers are collecting data *about* living individuals, or whether the focus is on the Internet censorship systems put in place. Regardless of the subject of investigation, the data collection procedure makes use of individuals' devices to initiate a measurement and collects their IP-address to verify the users' locations. The programme committee

makes the following statements, which indicate that they assume data is collected about living individuals:

1. "It is the PC's view that future guidelines should include as a core principle that researchers should not engage in experiments that *subject users* to an appreciable risk of substantial harm absent informed consent."
2. "[...] for Encore to work effectively at scale, *unsuspecting users* around the world must be enlisted to download oft-censored URLs without their informed consent."
3. "These requests could potentially result in severe harm; for example, when the *user* lives in a regime where due process for those seen as requesting censored content may not exist." (emphasis added, at top of Encore paper 2015 version)

Each statement indicates that living human Internet users are required, data about them is collected, and that the identification of these users through the collected data may cause them harm. I conclude that the programme committee considers Encore's data subjects as living individuals, and not just independent technical devices. This is relevant for an assessment through the lens of the ethics of care, because it can be argued that relationships between researchers and a multitude of data subjects are established through technical means.

Intervention and interaction

Intervention and *interaction* are concepts that stem from the biomedical centered approach of the Common Rule. These concepts can be applied to the data collection mechanism employed by Encore, too. The programme committee states:

"[...] to measure whether users in a given region like Switzerland can access URL X, Encore (deployed, say, at the Princeton web server) awaits an HTTP

request from a Swiss user and provides javascript that contains an embedded cross-origin request for URL X as part of the HTTP reply. The Swiss user's browser then executes the request for URL X, and the conditional javascript execution allows the Princeton web server to subsequently infer whether or not the request was censored." (p. 653, 2015 version)

The data collection mechanism *awaits* a request from a user, the user's browser then *executes* the request, and subsequently Encore's server *infers* the status of censorship. These italicised verbs all indicate that that some form of interaction is taking place between the researchers experimental design and the users, whether this interaction happens wittingly or not. The actions further indicate a newly established relationship. Further, the programme committee states that:

"[...] for Encore to work effectively at scale, unsuspecting users around the world must be enlisted to download oft-censored URLs without their informed consent." (p. 653, 2015 version)

The enlisting of unsuspecting users indicates, again, that an interaction takes place between the researchers and the Internet users.

Identifiable private information

The Encore research team also records the measurements, which includes the user's IP-address. Legal and policy interpretations in different jurisdictions appear to converge on the conclusion that an IP-addresses can identify an individual when relevant auxiliary data is available. With regards to Internet Service Providers (ISPs) the Court of Justice of the EU ruled in Breyer v. Germany in 2014 that an IP-address should be considered as *personal data* under the then applicable Data Protection Directive, when a party has the "legal means which enable it to identify the data subject with additional data which the internet service provider has about that

person.”⁹⁶ This position was reinforced in recital 30 of the General Data Protection Directive, which states that an IP-address, “when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them.”⁹⁷ Similarly the US Federal Trade Commission expressed that “persistent identifiers such as device identifiers, MAC addresses, static IP addresses, or cookies” meet the test of being personally identifiable information.⁹⁸ Courts in the US have offered a less clear legal position of the IP-address, though lawyers advise their clients to be cautious when collecting IP-addresses.⁹⁹

Further, when an unsuspecting Internet user’s browser sends a request to a potentially censored website, as instructed by the Encore code, their IP-address may be recorded by the server hosting the website, as well as by many intermediaries and potentially unknown third parties. Most significantly, government-mandated censorship systems may also record which IP-addresses attempt to access censored websites, although this is an assumption and the practice may vary significantly by country. It is likely, though, that a government actor can or has access to the required data to identify individuals on their national networks.

It can therefore be persuasively argued that Encore collects personal data (or ‘identifiable data’), or identifiable personal information. Further, Encore’s data collection mechanism may allow other parties to also collect the experimental data

⁹⁶ See

<<http://curia.europa.eu/juris/document/document.jsf?text=&docid=184668&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=1116945>>.

⁹⁷ See <<https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN>>.

⁹⁸ See <<https://www.ftc.gov/news-events/blogs/business-blog/2016/04/keeping-online-advertising-industry>>.

⁹⁹ See for example a blog titled “Are IP Addresses PII? Why Businesses Should Be Cautious About IP Addresses” by lawyers working for Fenwick & West LLP: <<https://www.lexology.com/library/detail.aspx?g=5648af5d-268c-4561-985c-9fcd2186c0a9>>.

and attribute URL requests to individual device owners. The authors of the Common Rule probably did not anticipate that a researcher's data collection actions may cause other parties—in Encore's case, the censor—to be able to collect data about the individual data subjects, too. It is uncertain how this would be considered in a strict interpretation of the Common Rule, though it is further analysed in the next part of this chapter. However, a researcher with a responsive attitude, would take care to understand the extent to which data is identifiable, or related to human subjects.

This section argues that the requirements for 'human subject' have been met, if the statements and assumptions of the programme committee are taken as a starting point for review. This establishes the study as 'human subject research' according to the Common Rule, so it is therefore subject to further review.

Before the review committee can assess whether the research is subject to any exemptions, article 46.101(e)&(f) require a legal assessment, stating that "Compliance with this policy requires compliance with pertinent federal laws or regulations which provide additional protections for human subjects" and "[t]his policy does not affect any state or local laws or regulations which may otherwise be applicable and which provide additional protections for human subjects." It is left unclear whether this legal assessment focuses exclusively on US law, or whether all jurisdictions of Internet users—potentially worldwide—needs to also be taken into account.

Interpreting "state or local laws" broadly, the Encore researchers would be faced with a global assessment of applicable laws for the protection of research subjects and data protection regulations. Further, the researchers would need to ascertain the extent to which their experimental design violates laws such as the US Computer

Fraud and Abuse Act¹⁰⁰ and similar laws around the world. This requirement is further discussed in the next part of this chapter.

Exemptions

No exemptions articulated in article 46.101(b) apply. The following paragraph briefly assesses each exemption:

- §46.101(b)(1): The research is not conducted in established or commonly accepted educational settings,
- §46.101(b)(2): The research does not involve educational tests, survey procedures, interview procedures, or observation of public behaviour,
- §46.101(b)(4): The research does not involve the collection of existing data, because the researchers developed their own data collection mechanism,
- §46.101(b)(5): The research does not study, evaluate, or examine public benefit or service programmes,
- §46.101(b)(6): The research does not involve taste and food quality evaluation and neither is it a consumer acceptance study.
- §46.101(i): The research does not explicitly involve prisoners or children.

Therefore, no exemptions of the Common rule would apply to the Encore paper.

No more than minimal risk

Article §46.110(b)(1) provides for an expedited review procedure of the study, if "some or all of the research appearing on the list and found by the reviewer(s) to involve no more than minimal risk". The programme committee makes some explicit

¹⁰⁰ See <<https://www.law.cornell.edu/uscode/text/18/1030>>.

statements about the potential for harm to unsuspecting Internet users whose devices are implicated in Encore's data collection mechanism:

"These requests could potentially result in severe harm; for example, when the user lives in a regime where due process for those seen as requesting censored content may not exist."

"At the same time, the risk and potential for personal harm associated with these measurements raise new challenges and ethical issues for our community that will warrant our close attention."

"It is the PC's view that future guidelines should include as a core principle that researchers should not engage in experiments that subject users to an appreciable risk of substantial harm absent informed consent." (p. 653 of 2015 version)

Given these statements, it is clear the programme committee determined that Encore does, in fact, pose potential *severe harm*, *personal harm*, and an *appreciable risk of substantial harm* to data subjects. Further, paragraph C of the OHRP Expedited Review Categories¹⁰¹ explicitly states that an expedited review procedure may not be applied where:

"[...] identification of the subjects and/or their responses would reasonably place them at risk of criminal or civil liability or be damaging to the subjects financial standing, employability, insurability, reputation, or be stigmatizing, unless reasonable and appropriate protections will be implemented so that risks related to invasion of privacy and breach of confidentiality are no greater than minimal."

¹⁰¹ This document can be considered to be an Annex to the Common Rule. See <https://www.hhs.gov/ohrp/regulations-and-policy/guidance/categories-of-research-expedited-review-procedure-1998/index.html>.

The Common Rule therefore requires that a full research ethics review of the study is carried out.

Waiver for informed consent

The authors did not seek the informed consent of their data subjects. The IRBs did not get to this stage of the research ethics evaluation, due to their reasoning about human subjects and minimal risk of harm. The programme committee, however, was clear that the study would present an “appreciable risk of substantial harm absent informed consent.” This section assesses whether informed consent could have been waived in this particular study. To reiterate the conditions under which a review body may waive the requirement for informed consent, §46.116(d) states the following:

An IRB may approve a consent procedure which does not include, or which alters, some or all of the elements of informed consent set forth in this section, or waive the requirements to obtain informed consent provided the IRB finds that:

1. The research involves no more than minimal risk to the subjects;
2. The waiver or alteration will not adversely affect the rights and welfare of the subjects;
3. The research could not practicably be carried out without the waiver or alteration; and
4. Whenever appropriate, the subjects will be provided with additional pertinent information after participation.

Assessing these criteria in turn (leaving out the last, as it pertains to deception research):

1. The programme committee considers the Encore study to pose potential severe harm, personal harm, and an appreciable risk of substantial harm, if conducted without the informed consent of implicated Internet users.
2. The programme committee states that some data subjects may indeed be violated in their rights or welfare, as many of the users will live under a regime where due process does not exist.
3. The Encore researchers make the following claim regarding the impracticality of carrying out the research with the informed consent requirement:

"For Encore, it would require appraising a user about nuanced technical concepts, such as cross-origin requests and how Web trackers work—and doing so across language barriers, no less. Such burdens would dramatically reduce the scale and scope of measurements, relegating us to the already extremely dangerous status quo of activists and researchers who put themselves into harm's way to study censorship [...]" (p. 664, 2015 version)

The authors consider informed consent to be a burden that will minimise the scale at which they can collect data. However, as described by (Dittrich et al., 2012), the exemption for impracticality presents a higher bar than claiming that obtaining consent is merely difficult or inconvenient. The burden expressed by the authors implies an inconvenience (ie. explaining how the data is collected to a lay person in a variety of languages). Displaying this information would be practically quite easy, for example by making the measurement widget visible (ie. increasing its size from "0") and displaying a message to the user, asking them to accept that measurements take place. In fact, displaying an informed consent notice is quite practical given the experimental set up of Encore.

Finally, the Common Rule provides two additional criteria that may allow a full or partial waiver for informed consent, neither of which is applicable to the Encore study. §46.177(c) asks whether (1) the consent documentation would be the only record linking the research and participant. Encore's system collects the IP-addresses, which are often considered as an identifiable data point. In section (2), the paragraph asks—again—whether the study would pose no more than minimal risk (which the programme committee does not consider to be true) and whether the study “involves no procedures for which written consent is normally required outside of the research context.” The Encore study makes use of a computer without the authorisation of its owner, which is a violation of the US Computer Fraud and Abuse Act (let alone all the other legal acts concerning the intrusion of computing devices worldwide). Authorisation, possibly in the form of informed consent procedure, would need to be obtained from the device's owner, so these criteria are not applicable to this case study.

Finally, the programme committee cites an independent survey study which concludes that respondents “agreed that most users for whom censorship is an issue would be unlikely to consent to Encore's measurements” (Schechter & Bravo-Lillo, 2014). This is a further indication that the requirement for informed consent should not be waived because of a weak impracticality argument.

Structured case research cycle analysis

Mirroring the structure of the previous chapter, this part the third research question for this case: “How might these concepts be reinterpreted to be useful and appropriate for the ethical assessment of technical Internet research?” This part continues to build the conceptual framework of the thesis.

The previous part identified the following concepts to be ambiguous in this case study:

- *Research*, specifically whether the Common Rule is applicable for Internet censorship measurement.
- *Human subjects*, specifically whether (1) the owners of devices implicated in Internet censorship research are data subjects, (2) whether the researchers interacted with data subjects, or somehow intervened, and (3) whether identifiable information was collected.
- *Informed consent*, specifically whether data subjects may experience harm, be violated in their rights, and whether obtaining informed consent is too impractical in this case.
- *No more than minimal risk of harm*, specifically the emergence of new challenges and ethical issues that arise from setting up downstream harms.
- *Privacy and legal assessments*, specifically the requirement for legal compliance with applicable laws of all data subjects and varying protections for human subjects in research around the world.

Research

The concept of *research* itself is not challenged in this case study. Following the concept's definition in §46.102(d), the Encore study is clearly a systematic investigation that aims to develop or contribute to generalisable knowledge. The IRBs did not refuse to review the study because it did not meet the criteria for 'research'. However, and interestingly, the authors and the program committee did claim explicitly that no set of rules exist to govern the research ethics review of a study like Encore. Burnett and Feamster write:

“Striking this balance between benefit and risk raises ethical questions that researchers in computer science rarely face and that conventional ethical standards do not address.” (p. 663, 2015 version)

And the programme committee mentions in their publication statement:

“The controversy arose in large part because the networking research community does not yet have widely accepted guidelines or rules for the ethics of experiments that measure online censorship.” (p. 653, 2015 version)

While the ethical issues raised around Internet censorship research and the Encore study specifically may not be common for computer scientists and review committees, does not mean that the Common Rule cannot be applied. The mock review in the previous part shows that the Common Rule would be able to impose requirements, such as informed consent, which the programme committees considered to be relevant and important. Specifically, the programme committee states that:

- “[...] researchers should not engage in experiments that subject users to an appreciable risk of substantial harm absent *informed consent*” (Publication statement, emphasis added).
- “Reviewers viewed such measurements as unethical and disavowed this Encore use case. While this ethical issue could be sidestepped by obtaining *informed consent* [...]” (Byers, 2015, p. 1) (emphasis added)

The statements from the authors and programme committee appear to be a reaction to the rejection for review by not one, but two IRBs. It is understandable (though alarming) that some computer scientists may not be fully informed about the concepts of the Common Rule, nor that IRB members are not fully aware of how technology and their users interact. However, their joint conclusion to not subject the

Encore study to a full research ethics review, rests on a misinterpretation: The IRBs did not consider Encore's measurements to include *human subjects*, as defined by the Common Rule. This concept is reviewed in the next section.

Human subjects

The interpretation of the concepts of the *human subject* is central to this case study. One IRB did not consider the data points to qualify as *human subjects*, whereas the other did not consider the issue in-depth. The latter's reasoning is that even if the data points would be human subjects in the sense of the Common Rule, they would not face any additional risk. Therefore, a full review would not yield more obligations for the researchers. This section first recalls how the concept of human subjects has been developed throughout this thesis. Then, the section assesses the facts, assumptions, and statements made with regards to the human subjects of the Encore in light of the constituent parts of the concept.

Development of the human subject concept

Chapter 2 presents the *human subjects* concept as relatively straightforward: a person is either a human subject in a research study, or they are not. The Common Rule stipulates that a human subject is a (1) living person, about (2) whom the researcher obtains data through intervention or interaction, which (3) is identifiable private information. This implies that the researcher knows to some extent who is in their research population and has some form of a relationship with them. This binary distinction is useful, since the classification of data subjects as human subjects triggers the research ethics review by an IRB.

Chapter 3 complicates the concept for social science Internet research in a variety of ways. First, computers and the Internet have disintermediated the traditional relationship between researcher and human (data) subject, creating a geographical and metaphorical distance between the two. Scholars have argued that IRBs are less likely to consider distant others to be considered as human subjects. Second, data subjects of Internet research may be located anywhere, because the global nature of the Internet overcomes previous physical limitations of data collection. The uncertainty and unfamiliarity of the data subjects' locations and their context makes it difficult or even impossible for the researcher to make a diligent risk assessment of the risks of research participation for an unknown data subject. Third, data about human subjects can be collected—through *interaction* or *intervention*—without the need for their permission, consent, or even knowledge. Scholars have stated that most research conducted on the Internet likely implicates humans in some way without their knowledge.

The case study in chapter 4 did not challenge, alter, or strengthen the concept of human subjects significantly. The password and username data used in the study was described as inherently private information, because it is the secrecy of this data that makes the most common authentication mechanism of the Internet function in the first place. This conforms to §46.102 (f)'s requirement for *identifiable private information* to have been "[...] provided for specific purposes by an individual and which the individual can reasonably expect will not be made public [...]". The paper in this case study does not, however, make use of existing data, so this particular criterium will not be developed further in this chapter.

Common Rule requirements

This section assesses the various requirements posed by the Common Rule to determine whether a research subject is, indeed, a human subject.

Living individual

The previous part of this chapter demonstrates that the programme committee—unlike the IRBs—considered the human owners of the implicated devices to be themselves enlisted in the study. In the programme committee’s public review, Byers repeatedly raises the relevance of the human owner of devices, their role as research participants, and the potential negative consequences they may need to endure:

- “[...] for Encore to work effectively at scale, unsuspecting users around the world must be enlisted [...]”
- “These requests could potentially result in severe harm; for example, when the user lives in a regime where due process for those seen as requesting censored content may not exist.”
- “[...] survey respondents of an independent study [...] agreed that most users for whom censorship is an issue would be unlikely to consent to Encore’s measurements.” (p. 1)

It is therefore reasonable to argue that researchers using the computing resources of data subjects in a way that may attribute actions to them, are interacting with living individuals and have created a relationship with living individuals. Afterall, Internet users need to actively visit a website that contains the Encore widget in its code, to (unknowingly) initiate a measurement for the research project, which implies that they are alive. This is further justified by the predominant interpretations of the legal

status of an IP-address, which states that an internet connected device is intimately linked to the person using or owning it, and third parties with auxiliary datasets may identify individuals based on unique data points such as IP-addresses. The data points may thereby qualify as living individuals and not mere technical artefacts. This position is affirmed and expanded on in the further sections.

Identifiable private information

The previous section of this part describes the dominant interpretation that IP-addresses can identify Internet users individually. The researchers collect the IP-addresses to identify the country or region within which the unsuspecting data subject initiated censorship measurements for the Encore project. However, the researchers themselves may not have access to relevant auxiliary data to identify their data subject individually. Therefore, the researchers may argue that they do not directly collect *identifiable private information*, though this claim can be contested.

The researchers may still be generating *identifiable private information*, by enabling third parties to collect data about the initiated measurements, too. Information about a device that triggered Encore's censorship measurements may give the appearance to a third party observing the URL request (such as a national censor or a local ISP), that a specific individual is requesting access to information which has been censored in a particular region. In the eyes of a third party, therefore, Encore's measurements attribute the actions of the research project to the owner of the device, who was implicated as a data subject without their knowledge. It is not inconceivable that a national censor or local ISP has access to the necessary auxiliary data (e.g. subscriber data or other unique identifiers for Internet access). It can be argued that the data collected by the researchers is indeed identifiable private

information that has been mistakenly attributed to a living individual. Therefore, the interpretation of the first IRB, who stated that Encore's data points did not qualify for the status of a human subject, can be argued to be incorrect.

Intervention/interaction (without permission/consent)

The Common Rule defines intervention and interaction with data subject as follows:

"Intervention includes both physical procedures by which data are gathered (for example, venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes. Interaction includes communication or interpersonal contact between investigator and subject. "

The researchers developed an intricate data collection mechanism that is invisible to the average Internet user. The user is therefore typically not aware that the researchers are communicating with their device, are initiating data requests, and thereby possibly manipulating their environment. Chapter 3 already noted that data about human subjects can be collected—through interaction or intervention—without the need for their permission, consent, or even knowledge.

The Common Rule is unclear about whether the human subject is required to have awareness of the interactions or interventions by the researchers. This awareness is likely assumed, since in the biomedical research realm for which the research ethics rules were initially written, the physician and patient would probably have a relationship through physical proximity and would be able to perceive each others' actions. This section argues that the researchers establish a relationship to human subjects through intervention and interaction, by making use of the affordances of the Internet and the reduced technical literacy of the average Internet user.

First, whether the interaction or interventions are visible or not, they can establish a relationship between the researcher and data subject. The data collection mechanism described in the Encore paper uses verbs that certainly imply an interaction:

Second, the researchers argue that an Internet user's claim to knowledge of how their devices are used by third parties in general has become diluted. They state (paraphrasing Princeton's IRB):

"It is generally accepted that users already have little control over or knowledge of much of the traffic that their Web browsers and devices generate (a point raised by Princeton's office of research integrity and assurance), which already gives users reasonable cover. By analogy, the prevalence of malware and third-party trackers itself lends credibility to the argument that a user cannot reasonably control the traffic that their devices send. The more widespread measurements like Encore become, the less risky they are for users." (p. 664, 2015 version)

The authors hereby equate their measurements to the multitude of third-party trackers and malware which inevitably control a part of the average user's device. The researchers likely try to argue that their intervention on the users' devices do not significantly manipulate the subjects' environments. This part argues that this is not necessarily true, since third parties may attribute sensitive actions to implicated human subjects. Researchers with a responsive attitude would seek to understand the change in environment their data collection mechanism effectuates.

The next section discusses that the resulting loss of autonomy may, in fact, incriminate unsuspecting Internet users and cause other harms. This raises

important questions, such as whether academic researchers should be held to a higher ethical standard than other data collectors on the Internet, and whether academic researchers should be permitted to manipulate the behaviour of Internet-connected devices to collect data about various aspects of the Internet.

Distance and the disintermediated research relationship

This part so far described what scholars have called the effects of the 'distance principle'. Aycock and Buchanan have described the phenomenon that as the distance between researchers and data subjects increases, the former are less likely to consider their data points as affecting humans, and therefore are less likely to consider their research as involving human subjects (Aycock, Buchanan, Dexter, & Dittrich, 2011; Buchanan et al., 2011). The Internet has afforded the ability to collect data anywhere in the world for zero marginal cost, which has disintermediated relationships between researchers and data subject further, whereby the experiential distance has decreased the awareness that a research relationship still exists (Frankel, M. S. & Siang, 1999; Keller, H. E. & Lee, 2003).

The statement by Georgia Tech's IRB illustrates the loss of perceived personhood by a distant data subject also for research ethics review boards. In their email to the researchers, the IRB states "this does not require IRB review since it doesn't meet the definition of a human subject." It may be in the short-term interest of computer scientists to present their data subjects as mere technical data points to avoid the imposition of responsibilities, duties, and a delay of their data collection. It is worrying, however, when an IRB does not challenge this position, when a relatively simple sociotechnical analysis reveals that the use of devices for data collection may attribute incriminating actions to implicated humans.

In this case study, the researchers' and their IRBs' judgement was affected by the distance principle, which led to a dehumanisation of their data subjects. In her paper reflecting on the research ethics process at the University of Ontario, Dominique Rivière describes that an IRB cannot always prevent the dehumanisation of research subjects in the course of a study. However, she argues the IRB "[...] can—and, in my view, should—foreground for researchers that they may be confronted by such possibilities, and ask them to make educated guesses about how these possibilities might be handled" (Rivière, 2011, p. 203). She advocates for a research ethics review process that supports researchers as they deliberate "[...] on the ethical dilemmas, tensions and issues that are specific to the nature and context of the research they are conducting [...]". In making her argument, she leans on Magda Lewis, who argued that it is not necessarily the research analysis that is considered unethical, but instead the failure to critically assess the research's "social/political engagements in a human context" (Lewis, 2008, p. 693). Following Rivière and Lewis, research ethics review board such as an IRB therefore has a duty to assist researchers to overcome the distance principle and to understand the human context within which their interventions and interactions take place.

Informed consent

The researchers challenge the requirement for informed consent in their paper with three separate arguments. They argue that (1) assessing the benefits and risks of a study like Encore's is difficult, (2) obtaining informed consent would, in fact, place research subjects in a position of more potential harm and complication of their rights, (3) that it is impractical to inform potential research subjects of the risks in a

language they can meaningfully comprehend, both in terms of technical terminology and vocabulary.

(1) Benefits and harm

The researchers articulate the benefit of a research study like Encore clearly: "Encore enables researchers to collect new data about filtering from a diversity of vantage points that were previously prohibitively expensive to obtain and coordinate.

Ongoing efforts to measure Web filtering would benefit from Encore's diversity and systematic rigor [...]". The authors also incorporated much of the criticism on their first paper, and articulated their understanding of the risks posed by Encore as "nebulous". The laws and political sensitivities of accessing the websites targeted by Encore indeed vary across the world, resulting in "no ground truth about the legal and safety risks posed by collecting network measurements." However, discarding the legal and safety risks to research subjects because they are not easily quantifiable, is not a reasonable approach. In response, the researchers did limit Encore's URL requests to a few less controversial websites, albeit somewhat reluctantly.

(2) Rights and welfare

The researchers argue that obtaining informed consent from data subjects would potentially increase the risk to data subjects, as it would remove their claim to plausible deniability. This argument is understandable in comparison to how censorship was previously measured with people on the ground actively deploying measurement systems. However, the researchers thereby apply the assumptions of the US justice system where plausible deniability may be a legitimate defence, to other jurisdictions where no guarantee of a fair trial by an independent judiciary may exist. It is not unreasonable to expect countries that censor much information to

also exercise their power over citizens in other ways, such as in politically motivated trials. Finally, research subjects may be framed for espionage, since Encore collects operational data about critical information infrastructure and sends it back to the US. It is short-sighted, selfish, and reckless to assume that these measurements can not cause harm and that unsuspecting research subjects have a legal defense. Rights and welfare therefore need to be understood and interpreted in the context of data subjects, rather than from the jurisdiction the researchers are most familiar with.

(3) Impracticality

In a further attempt to circumvent the requirement for informed consent, the researchers state that fulfilling this requirement would be impractical as it would entail teaching users “nuanced technical concepts... across language barriers.” The challenges of communicating the necessary technical information to a global set of participants again highlights the tension between the scalability imperative in computer science and respecting the expectations of informational norms.

Interestingly, they further note that informing users would “dramatically reduce the scale and scope of measurements”, which is in part based on a survey about whether engaging in particular forms of research is considered to be 'ethical', as conducted by Stuart Schechter and Cristian Bravo-Lillo. These authors report that "the two experiments that our respondents objected to the most were the censorship-detection studies from the networking measurement community" (Schechter & Bravo-Lillo, 2014). This implies that, if Internet users would learn that their devices may be used to collect data about Internet censorship methodologies, they would be most likely to refuse to participate. This should be an indication to researchers and IRBs with responsive attitudes that the covert execution of such research is unlikely to be a well-received idea.

Autonomy

Chapter 2 presents autonomy as a fundamental value in modern Western philosophy, which underpins widely supported principles such as the social contract and popular sovereignty. These are considered to enable rational choices of individuals and the citizenry as a whole, whereby the institutions of political power are legitimised to be instrumental in the achievement of citizen welfare. Kant positioned autonomy to be the key to understanding the moral obligations that people have towards each other. John Christman summarizes John Stuart Mill's conception of autonomy as "the capacity to be one's own person, to live one's life according to reasons and motives that are taken as one's own and not the product of manipulative or distorting external forces" (Christman, 2008).¹⁰² Autonomy therefore underpins other fundamental values, such as freedom and liberty.

For studies such as Encore, Philip Pettit's theory of 'Liberty as Non-Domination' offers a useful starting point to assess whether the researchers' intervention and interaction (or: interference) respect the autonomy, and therefore the liberty, of Internet users. According to Pettit, liberty is achieved when actors in society do not arbitrarily dominate each other. In his words:

"Interference involves an intentional or quasi-intentional worsening of someone's choice situation and occurs on an arbitrary basis to the extent that it is not forced to track the interests and ideas of those who suffer the interference. One party dominates another just so far as they have the capacity to interfere on an arbitrary basis in some of the other's choices" (Pettit, 1997, p. 1)

¹⁰² Online publication. Available at: <https://plato.stanford.edu/archives/spr2018/entries/autonomy-moral/>.

Applying this theory to the case, researchers must be able to perform the following to violate the autonomy and liberty of Internet users:

1. Have the capacity to interfere,
2. On an arbitrary basis,
3. To (quasi-)intentionally worsen someone's choice situation.

Pettit's theory of Liberty as Non-Domination applies seamlessly to a project like Encore. This section addresses some of the statements of the researchers as described in the previous section, through the lens of Pettit's theory. The section ends with a description of what a dominating relationship between researcher and research subject may mean for the researcher's relationships to society and to their colleagues.

1. Capacity to interfere

The researchers developed a mechanism that initiates URL requests from an unsuspecting person's device. The section above titled 'Intervention/interaction (without permission/consent)' concludes that the data collection mechanism intervenes and interacts with data subjects. These words are synonymous or have a similar meaning to the word *interfere*. The researchers therefore have the capacity to interfere with unsuspecting Internet users, which further establishes a relationship with them through technical means.

2. Arbitrary basis

Measurements are initiated from devices used by people who access specific websites. These websites do not inform their visitors that experimental measurements are conducted upon loading and rendering the website's code. The researchers have little influence over who visits the website, beyond targeting

websites that have an audience in specific regions with a request to include the code in their website. Therefore, the interferences are relatively arbitrary. It may be argued that the established relationship is therefore with society as opposed to specific, individual Internet users.

3. Worsening choice situation

The unsuspecting research participants are unaware of the measurements that are conducted using their Internet connection, and neither are they able to stop them from happening. They do not have a choice in the act of requesting websites that are potentially considered to be controversial in their sociopolitical context. Further, they are not aware the data collection mechanism may be attributing actions to them, in the eyes of other, possibly more powerful actors in their context. Therefore, their choice situation as to whether to access these censored websites is worsened (or even non-existent).

Given Pettit's description of domination, it is clear that the researchers dominate their research subjects, which violates their liberty and therefore their autonomy. The extent of domination extends further to all Internet users, however. Not only could anyone—in theory—access a website that has the Encore widget installed.

Further, by publishing this paper, the authors are incentivising their colleagues to use similar means to collect data at a distance. Had it not been for the stern warning in the submission statement of the program committee, the publication of this paper would have set a precedent for researchers to make use of unsuspecting Internet users' devices for data collection. It may still be argued that this paper sets a precedent, regardless of the submission statement, because it is an efficient

methodology to collect data from vantage points around the world. This dominating role of researchers over the rest of society must be seen as a negative relationship, which is not only characterised by domination, but also selfishness, recklessness, and exploitation of a technically ill-informed, and therefore vulnerable population. This analysis is relevant for the section on moral wrong and moral harms.

No more than minimal risk of harm

The concept of 'no more than minimal harm' has been developed thoroughly in this thesis so far. Chapter 2 describes the Common Rule formulation, which consists of a 'probability and magnitude of harm' calculation, a standard of harm being "not greater in and of themselves than those ordinarily encountered in daily life". The chapter observes that the concept appears to have been developed specifically for clinical research, because it takes the baseline of "[...] the performance of routine physical or psychological examinations or tests" to apply the latter standard in research ethics review.

Chapter 3 notes that the application of this concept is complicated for Internet research, because the data subjects of a particular study may be dispersed around the world, instead of in a confined clinical setting. Therefore, their physical and sociopolitical context are likely to differ, potentially significantly. The relevant variables in a calculation of the probability and magnitude of harm or a 'daily life'-standard therefore become increasingly divergent, whereby it is impossible and irresponsible to apply a single standard to all data subjects.

Chapter 4 constructs some clarity on the concept of 'harm' and positions it in the realm of academic experimentation. The chapter describes that harm can be inflicted on a persons' interests (e.g. financial, physical), that it may be a wrongful act that does not damage the persons' interests necessarily, or it could be understood as a moral harm that is to be tolerated though may violate values such as someone's dignity. Further, the chapter differentiates between the harms to data subjects, other researchers or the research profession, and society as a whole.

This section builds on the development of this concept so far and consolidates the section regarding privacy. The authors identify a number of potential harms in their paper, though describe them as being “nebulous”:

“The risk that Encore poses are far more nebulous: laws against accessing filtered content vary from country to country, and may be effectively unenforceable given the ease with which sites (like Encore) can request cross-origin resources without consent; there is no ground truth about the legal and safety risks posed by collecting network measurements.” (p. 663, 2015 version)

These harms, along with a few others identified in this section, are interlinked with an impact of the research subjects' privacy and a violation of numerous other rights or interests. For example, other harms include a violation of research subjects' privacy rights or rights to data protection laws, violation of computer intrusion laws, defying the reasonable expectations of Internet users about how their devices and data are used, attributing actions research subjects, and violation of other relevant laws (such as accessing censored URLs) which may vary in different jurisdictions. In the first iteration of the Encore paper, the authors state that while not all effects of the Encore project are unlawful, the resulting consequences could still be harmful:

“Our conversations with citizens of various countries suggest that attempting to access censored content is not illegal per se; nevertheless, simply because an activity is legal does not make it harmless, so we are sensitive to inducing users to retrieve content that might land them in trouble.” (p. 12, 2014 version)

Specifically, this section focuses on the *intrinsic* harms to the rights and welfare of data subjects as being participants in an experiment, and the *extrinsic* harms by third parties to the data subjects, to the research profession, and to society. Ben Bradley describes the difference between intrinsic and extrinsic harm as follows:

"Some events are intrinsically harmful: their mere occurrence constitutes harm to the person to whom it is occurring. Others are extrinsically harmful: they are harmful because of their effects." (p. 392)

Finally, the section compares the foreseeable harm to those that have given rise to the related legal doctrine called 'negligence' or the 'duty of care'. First, however, the various actors who may cause harm are described.

Actors and attackers

To assess potential harm and its magnitude resulting from the Encore study, it is useful to briefly reflect on which actors may cause harm, and to whom. This section divides the actors into active, passive, and active third parties.

Active

Two active groups exist in this study, who make active, deliberate decisions. First, the researchers developed Encore's data collection mechanism to actively initiate measurements from the devices of Internet users. Second, the data collection mechanism was implemented by owners of (or other persons responsible for)

particular websites. These actors are unlikely to have malevolent intentions with the welfare of implicated data subjects.

Passive

Two passive groups are involved in this study. First, the Internet users whose devices are implicated in Encore's data collection mechanism do not initiate the measurements themselves. Second, the owners of intermediary Internet-connected systems (such as the owners of servers, ISPs, and those operating the censorship mechanisms) route the data requests from the Internet users' devices to the requested URL or censorship mechanism, and the response from the devices to the research server in the US. This second group may overlap with the next group.

Active third-party

The measurements initiated on Internet users' devices by Encore's data collection mechanism generate a series of data flows that are routed and possibly recorded by third parties. If these parties do not act upon this information, they remain in the 'passive' group above. However, given a combination of capacity, potential political expediency, or general enforcement orders, these third parties who take note of the information flows may decide to act. For example, a censor in an authoritarian regime may have decided to arrest persons who attempt to access controversial content that has been censored in their jurisdiction. They may collect IP-addresses of the devices that activated the censorship mechanisms, or compel intermediaries to inform the authorities when a device appears to be requesting specific URLs. The active group thereby enables active third parties to respond to the requests for censored information. Their actions may result in extrinsic, or downstream harms.

Intrinsic harms

Following the description by Ben Bradley, intrinsic harms occur to participants simply by being implicated in Encore's data collection mechanism. This section builds on the findings of chapter 4 to assess the impacts of the Encore study on research subjects' interests, wrongs inflicted on them, and resulting moral harms. The harms to society and the research profession are a consequence of the Encore study, but do not result intrinsically from the study. They are therefore discussed in the next section on 'extrinsic harms.'

Interests

Chapter 4 describes a harm to a person's interests as the type of injury that may be compensated for through legal procedures. This implies that these harms are covered by existing law in one way or another. The Common Rule even states explicitly in §46.101(e, f, g) that researchers must comply with relevant and applicable federal, state, local, and foreign laws.

Data subjects implicated in Encore's data collection mechanism may be violated in several rights: First, their data protection or information privacy rights may be violated if their IP-addresses are collected—without their informed consent—along with other identifying, auxiliary information. Further, owners of the implicated devices may invoke computer intrusion laws, which exist in different variations and areas of legal practice in jurisdictions around the world. In essence, computer intrusion laws determine it to be unlawful to use or interfere with the computers owned by other parties without their explicit consent.

The implicated research population of the Encore project is potentially spread over many jurisdictions in the world. The requirement in §46.101(g) to comply with applicable foreign laws therefore complicates the legal diligence required by responsive and context-aware researchers. The researchers stated themselves, aptly, that applicable “laws [...] vary from country to country”. The programme committee further alerts to a variation in principles and assumptions in legal systems when they discuss the consequences of research “[...] in a regime where due process for those seen as requesting censored content may not exist.”

It will be difficult, if not impossible, for a technical research team to carry out a global survey of applicable laws and policies that may harm their dispersed research population. If a worldwide legal study in, indeed, deemed to be infeasible, technical Internet researchers cannot guarantee that their experiments and data collection mechanism do not violate any laws. The ethical considerations of putting a data subject into a position where a third party may conclude they are violating a law is discussed in the next section. On the intrinsic violation of laws, however, the Menlo Report advises researchers to consider whether a violation of laws is justifiable when it states:

“Respect for public interest can often be addressed by obeying relevant laws. If applicable laws conflict with each other or with the public interest, and a decision is made to not comply with legal obligations that are viewed as unethical, researchers should have ethically defensible justification and be prepared to accept responsibility for their actions and consequences.” (p. 14)

Moral wrongs and moral harms

Data subject’s welfare may also be violated, beyond violations of their rights. For example, the data collection mechanism may violate their expectations of how their

devices work, may violate their autonomy and dignity, and may implicate a vulnerable group of unsuspecting persons. A larger range of conceivable moral wrongs and harms exist, though this section is limited to these three.

For the intrinsic harms, the most relevant is the violation of the expectation of how implicated research subjects' devices work. Encore researchers justify the functioning of their data collection mechanism with those used by the advertising industry. A variety of studies have found, however, that the underlying mechanisms for online advertising and behavioural marketing cause discomfort for the average Internet user, when presented with their actual functionality. Internet users' expectations of tracking and the use of their devices by third parties has been found to be at odds with technical reality (McDonald & Cranor, 2010; Turow, King, Hoofnagle, Bleakley, & Hennessy, 2009; Ur, Leon, Cranor, Shay, & Wang, 2012). The research ethics guidelines published by the Association of Internet Researchers (AoIR) had foreseen this issue. They ask their community: "What are the ethical expectations users attach to the venue in which they are interacting, particularly around issues of privacy?"¹⁰³ While commercial companies may have set precedents that Internet users' devices trigger more requests than they expect, the AoIR guidelines attempt to hold academic researchers to a higher ethical standard.

Similarly, a responsive researcher in the sense of the ethics of care would be expected to take into account this type of research when interpreting their data collection mechanism through analogy to another type of technology. Without further consideration of the analogy, the researcher would not demonstrate an

¹⁰³ See p. 8, accessible at <https://aoir.org/reports/ethics2.pdf>.

appropriate identification, understanding, and interpretation of the negative outcomes for data subjects.

Extrinsic harms

Extrinsic harms are those that occur as a consequence of an action but are not directly inherent to the action. They occur 'downstream' from the action in a causal chain that is initiated by the agent. Much literature exists in law, philosophy, and ethics about the extent to which an agent can be held morally or legally responsible for an action they have caused. A key factor in determining responsibility, is the extent to which the agent acted with the intention to cause an outcome, or whether the outcome is the result of negligent behaviour.

The interests of implicated Internet users as research subjects are described in the section on intrinsic harms. This section builds on the loss of autonomy by research subjects described in a previous section, and its resulting moral wrongs and harms to data subjects.

In the case of Encore, extrinsic harms may occur as a result of actions by third parties who act upon the information flows initiated from an unsuspecting Internet user's device. Such harms can be plentiful, since the measurements may be triggered on devices that are located on Internet networks in a variety of geographic and thus socio-political environments. For example (without singling out particular countries), a range of jurisdictions censor websites, track their citizens' online behaviour, and are known to arrest people based on their online activity. It is therefore not inconceivable that a national censor can keep a record of who in their

jurisdiction has attempted to access censored websites.¹⁰⁴ Further, Internet censorship is a more common phenomenon in authoritarian regimes, where rules consider the free flow of information and ideas to be a threat to their position of power.¹⁰⁵ Authoritarian regimes are more likely to abuse the rule of law for political ends, whereby Western concepts of a 'fair trial' or 'innocent until proven guilty' are not safeguards citizens can count on (Henderson, 1990). The programme committee of the ACM Sigcomm conference noted that the resulting harms by third parties can be serious:

"These requests could potentially result in severe harm; for example, when the user lives in a regime where due process for those seen as requesting censored content may not exist." (p. 1)

More concretely, the loss of agency over their device, and thus worsening of the choice situation of an implicated research subject, means that a third party such as a censor may attribute actions to individuals. For example, the measurements may be interpreted by a censor as support for a cause, such as those that are deemed inappropriate within the jurisdiction in which it is censored. Further, if the government's Internet surveillance agencies can trace that a citizen's device is transmitting data regarding its network, and especially its national censorship system, it could suspect this person is engaged in espionage. Depending on the investigative and enforcement measures, it is thus conceivable that the Encore researchers are setting up at least some of their unsuspecting research subjects for severe harm, even though it would be caused by others. It is clear from this analysis that the implicated Internet users in Encore's data collection mechanism are

¹⁰⁴ Political arrests for politically motivated online activity are well documented in Freedom House's website and annual report: <<https://www.freedomonthenet.org/countries-in-detail>>.

¹⁰⁵ See <<https://firstmonday.org/article/view/876/785>> and <<https://www.hoover.org/research/authoritarianism-vs-internet>>.

therefore a vulnerable population. A researcher with a responsive attitude would need to identify and understand these harms, and interpret harm mitigation in their experimental set-up.

Degrees of responsibility for downstream harms

Responsibility for the consequences of research are limited in a variety of ways. This section explores the limitations provided in the Common Rule, in general ethics literature, and in law. The aim of this assessment is not to conclusively determine whether the Encore researchers should be held responsible for hypothetical enforcement actions, but to develop a more general understanding of the considerations that may contribute to a determination of whether technical researchers should be held responsible for the information flows they initiate.

In §46.111(2), the Common Rule states the following:

Risks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result. In evaluating risks and benefits, the IRB should consider only those risks and benefits that may result from the research (as distinguished from risks and benefits of therapies subjects would receive even if not participating in the research). *The IRB should not consider possible long-range effects* of applying knowledge gained in the research (for example, the possible effects of the research on public policy) as among those research risks that fall within the purview of its responsibility. (emphasis added)

This provision requires IRBs to limit their assessment of the consequences of research based on the following criteria:

1. Assess whether the risks to research subjects are reasonable in relation to the anticipated benefits to:
 - a. the research subjects, and,
 - b. the importance of the knowledge that may reasonably be expected to result;
2. Only consider the risks and benefits that may result from the research;
3. Disregard the possible long-range effects of applying knowledge gained in the research.

This limitation does not differentiate between intrinsic and extrinsic harms, though it does exclude consideration of the longer-term effects of third parties applying the knowledge. The focus is on applying the research results but does not exclude responsibility for creating a causal sequence, whereby potentially incriminating actions are attributed to unsuspecting data subjects, which in turn may be misinterpreted by powerful third parties. The extrinsic harms discussed in this part would, in principle, fall within the scope of the Common Rule.

Some current debates in ethics literature give some further indication of how an IRB may interpret the extent of responsibility for technical researchers. Michael S. Moore assessed the moral underpinnings of attributing responsibility for enabling harmful consequences by third parties in his book 'Causation and Responsibility: An Essay in Law, Morals, and Metaphysics' (Moore, 2009). In essence, Moore argues that an agent can still be held responsible for acts when they do not directly cause a harmful consequence, but initiate a causal sequence of events that would not have existed without their (negligent) act. He proposes the *counterfactual dependence* test, which assesses in the counterfactual whether an outcome would have come to bear, if it

were not for the agent's initial action. He discusses the *substantial cause requirement* to assess whether the degree of causation of the (negligent) act was, indeed, substantial enough to justify attributing responsibility to the actor. Gideon Rosen responds critically to Moore's work, arguing that it does not necessarily contribute new insights beyond current approaches in law, however (Rosen, 2010).

Legal approaches to extrinsic harm

Legal systems worldwide assign liability for downstream harms in a variety of related ways. In general, the legal mechanism to ascribe responsibility to an actor is the tort law of *negligence*. This legal principle does not necessarily demand careful conduct directly, though it places responsibility on an actor when their lack of care results in an injury to a third person. It thereby aims to encourage the exercise of reasonable care by actors, taking into account the potential harm they may cause. Negligence and its related concepts can be assessed in a variety of ways. This section discusses two examples as an illustration of legal tools and their requirements to assess whether an actor may be held responsible for an action that leads to extrinsic harms. It is beyond the scope of this thesis to provide a full overview of the numerous court cases and legal literature.

Proximate Cause and Sine Qua Non

The proximate cause doctrine has been developing in US jurisprudence and legal research as one of several methods to ascribe causation to an actor, even if their actions did not directly cause a harmful event. Patrick Kelley explains in his paper 'Proximate Cause in Negligence Law: History, Theory, and the Present Darkness' that the proximate cause doctrine has been difficult to define precisely, and that looking at leading case law, "we find only puzzles and confusion" (Kelley, 1991, p.

51). Sandra Sperino clarifies that courts have not arrived at a consistent set of concerns that underlie proximate cause, because the concept continues to evolve (Sperino, 2013). In short, however, a court of law can reason that an act is sufficiently related to an injury that the action may be considered to be the cause of the injury. Judges apply a "but-for" test, also known as *sine qua non*, to assess whether the injury would have occurred without the agent's action. Sperino states that the proximate cause concept allows judges and policymakers to express a normative preference about where the line for liability regarding downstream harms should be drawn.

Professional negligence and a special duty of care

Legal systems typically demand a higher standard of care from professionals who maintain a level of expertise or skill. For example, the remedial legal constructs such as legal malpractice or medical negligence are designed to ensure society can place their trust in lawyers and doctors, knowing that their best interests are served. These constructs impose a special *duty of care* on professionals, which is based on their superior knowledge and levels of trust required to fulfil their roles in society. The expected duties imposed on professionals are higher than the typical care expected from a 'reasonable person', because they present themselves as having more knowledge or skill than the average person. Duties of care are also imposed on other professions, such as manufacturers, architects, and accountants.¹⁰⁶

While a clear duty of care may not have been imposed on the research profession, the various articulations of research ethics procedures and concepts fulfil this role. A duty of care for researchers with regards to extrinsic harms, may be further informed

¹⁰⁶ See, for example for manufacturers <https://www.jstor.org/stable/1285619?seq=1#page_scan_tab_contents>, for architects <https://link.springer.com/chapter/10.1007%2F978-1-349-13214-0_6>, and for accountants <https://www.accaglobal.com/content/dam/acca/global/PDF-members/2012/2012p/Prof_liability.pdf>.

by professional negligence and a fiduciary duty of care. It is beyond the scope of this thesis to offer a review of the court cases and literature regarding these two concepts, but a high-level overview is relevant for the interpretation of the responsibilities for the social position academics hold.

Professional Negligence

English law offers a well-developed set of criteria to determine whether a professional acted negligently in the performance of their services. First, the defendant must have presented themselves as a professional in a particular field, with the knowledge or skill that the claimant sought. This establishes a duty of care for the agreement between the defendant and claimant. Second, the duty of care was breached by the defendant, for example by advising an action when, in fact, this does not meet the standards one would expect from a professional in their position. Finally, the claimant needs to demonstrate that they have been injured or suffered some form of loss as a result of the negligence by the professional.¹⁰⁷

This doctrine is not directly applicable to researchers who make use of the devices owned by unsuspecting Internet users for their research projects. However, the parallels are clear: researchers, like doctors or lawyers, hold a powerful position in society which allows them to conduct studies and collect data in the pursuit of knowledge. The ethics of care may provide the tools for an interpretation: First, the relationship established through technical means between the relatively powerful researcher and the data subject is unknown to the latter party. The researcher presented themselves invisibly. Second, the researcher is able to alter the device of the data subject, without due consideration of potential harms. Third, if a harm

¹⁰⁷ [1963] 1 Lloyd's Rep 485, [1963] 2 All ER 575, [1964] AC 465, [1963] 3 WLR 101, [1963] UKHL 4, available at: <<https://www.bailii.org/uk/cases/UKHL/1963/4.html>>.

would occur to the data subject, the researcher would be held accountable based on professional negligence law. A responsive researcher would identify, understand, and interpret the relationship, and take upon themselves the duty to mitigate any foreseeable harm.

Special fiduciary Duty of Care

Dutch banking law prescribes a fiduciary duty of care for banks and credit institutions to protect consumers against irresponsible financial risks. Case law extends this duty also to third parties, who may be implicated in irresponsible decisions made by a bank, but do not have a contractual relationship with the bank. In the case MeesPierson/Ten Bos, the Dutch Supreme Court articulated a new rule, stating that—because of their powerful position in society—banks enjoy a *special* duty of care. In this case, a mother had added the names of her adolescent daughters to a savings account. A representative at the bank had invested their funds in speculative ways, leading to a significant loss of wealth. Speculative investments are unlawful with the accounts of adolescents in the Netherlands, however. The special duty of care imposed on banks requires investors to assess which parties will be affected by potentially speculative and thus somewhat irresponsible decisions.

The interesting aspect of this doctrine for this thesis is the requirements for powerful actors in society to actively take into account the effects of their actions on implicated parties beyond the first point of contact. Still, the special duty of care can only be called upon when a claimant has, indeed, suffered an injury or loss of some sort. It does pose an interesting line of thought for the duty of care of technical Internet researchers, who may set up harmful situations for implicated third parties.

Conceptual Framework: 3rd Iteration

Human subjects	
Relational approach to morality	Data subjects: • Technological mediation ◦ Relationship established through the use of networked technology, ■ <u>Technical mechanism initiates data flows on device of unsuspecting data subjects.</u> ■ <u>Internet connected device is typically intimately linked to person owning it.</u> ■ <u>Data collection depends on living persons undertaking an action (e.g. visiting a website), it means unsuspecting Internet users are enlisted into the study.</u> ■ <u>Process of data collection (intervention/interaction) indicates a <i>dependency</i> relationship between researchers and data subjects.</u> ■ <u>Potentially attributing actions to users establishes a XXX relationship.</u> ◦ Enrolment into relationship may not be voluntary, ■ Consent technically not necessary for interaction or intervention, ■ <u>Knowledge on the part of data subject is irrelevant,</u> ■ <u>Average data subject has limited knowledge of information flows on their devices in the first place.</u> ◦ <u>Technical data points may qualify as data about human subjects</u> ■ <u>Using affordances of the sociotechnical network to collect data.</u> ■ <u>IP-address can be considered to be identifiable information.</u> • Distance principle ◦ Disintermediated relationship through affordances of networked technology, ◦ Increased experiential distance of researcher ◦ Researchers and data subjects not easily identifiable ◦ Data subject is <i>distant other</i>. • Passive data collection, ◦ Received from third party, ◦ No intervention or interaction. Society: • Hierarchical societal position of researcher ◦ Academics enjoy a powerful position, ◦ Technologists shape human communications and interaction, • Random people can be implicated in an Internet research study by virtue of using the Internet. ◦ <u>Data collection mechanisms may make use of affordances of the Internet to observe behaviour of technical components without the knowledge of their owners/users.</u> • Power relationship to data subject, ◦ Potentially oppressive relationship, ◦ Data subject is vulnerable to the decisions made by researchers, ◦ <u>Researchers have ability to attribute actions to actors without their knowledge by making use of their devices and Internet connections.</u> • Data subject's data is technically publicly available, ◦ Not publicly available in the meaning of the concept. Research profession: • <u>Academic profession to be held to a higher standard than other technical actors?</u>
Contextually-oriented deliberation	Research context: • Assessing the implications of a disintermediated relationship is based on a lack of knowledge, • <u>Active data collection mechanism may generate identifiable data by enabling active third parties with auxiliary data to re-identify Internet users.</u> • <u>Consider assessing the 'human subject' concept on a scale risk of harm potential.</u> Data subject context: • Difficulty or impossibility to know local context of data subjects, • <u>Measurements may alter the behaviour of devices.</u> ◦ <u>Which may be the catalyst to change the data subject's context,</u> ◦ <u>Which may inadvertently cause harm to the data subject (e.g. incriminate Internet users in activity in which they played no intentional part).</u>
Responsiveness	Identify: • Human ownership or use of implicated devices,

	• <u>Create awareness of whether and how researchers change the data subjects' environments.</u> • <u>Create awareness of whether data subjects are being dehumanised as mere data points.</u> Understand: • Data collection through intervention or interaction with devices may implicate humans, • <u>Need to surmount language and cultural barriers to communicate information about data collection and research effectively.</u> • <u>Existing surveys that uncover the perceptions and preferences of Internet users should inform decision-making about the appropriateness of those data collection mechanisms.</u> Interpret: • Sensitivity to the offline context within which the data subject is likely nested.
Emotions as moral tools	Positive: • Beneficial research, ◦ By benign researchers. • Possible to understand and meet the expectations of Internet users. Negative: • Humans may be used as a means to an end, ◦ If local context is not identified and understood, ◦ And researcher is not responsive to their specific needs. • Research can be considered as a form of surveillance

Minimal risk of harm	
Relational approach to morality	Data subjects: • <u>Research ethics gatekeepers appear to accept that users have little knowledge of web traffic generated by their devices.</u> • Harm can materialise in a variety of ways ○ Interests ○ Wrongs ○ Moral harms • <u>Researchers maintain an active connection through website, devices, and servers.</u> ○ <u>Further connection enabled to servers hosting potential censored content.</u> ○ <u>Which in turn connects to active third parties (e.g. government censors).</u> • <u>Attributing potentially incriminating actions to data subjects establishes a (special) relationship.</u> Society: • Weigh harms against the social importance of the research • Injury to society ○ Committing harms sets precedent for: ■ Inappropriate acts by researchers, ■ Inappropriate acts by malevolent parties. • Effect of harms to relationship with society, ○ Knowledge gain to research disciplines, vs, ○ Higher-order loss of trust to research discipline. Research profession: • <u>Creating incentives for colleagues to make use of third-party devices in inappropriate ways</u> • Loss of public trust in research discipline may affect availability of funding. Active third parties: • Researchers enable third-parties to collect data and act, ○ creating downstream harms. • Third parties may act upon information flows, depending on: ○ capacity, ○ potential political expediency, ○ enforcement orders.
Contextually-oriented deliberation	Research context: • <u>Potential risk to complex web of stakeholders</u> ○ A single 'daily life' standard does not exist, ■ Variables in the typical calculation of the magnitude of risk multiplied by its probability are too complex and vague to be meaningful, ○ Legal diligence may need to be global in nature, ■ <u>Variations in the principles and assumptions, due process may not be guaranteed.</u> ■ <u>Researchers cannot guarantee that research will not violate any laws.</u> ■ <u>Violation of laws need to be defensible, researchers may need to accept responsibility for resulting harm.</u> • Researchers may not be able to assess whether their actions constitute a minimal risk of harm. • Vulnerability: ○ Data subjects may be vulnerable persons according to existing definitions, ○ Data subjects may be vulnerable as a result of being implicated in the research. Data subject context: • Harms are highly contextual, • Data subjects can be situated in wide-ranging contexts with undefined sociopolitical parameters, • Research may create new risks of harm in the data subject's context: ○ By disregarding local (informational) sensitivities, ○ By being insensitive or ignorant to existing power relations, ○ By violating local laws, ○ <u>Attributed actions may be incriminating in some jurisdictions (where legal principles may vary, to the detriment of Internet users).</u> ○ <u>Enabling extrinsic (downstream) harms.</u> • Violations of moral harms include: ○ autonomy ○ privacy ○ personal integrity ○ dignity
Responsiveness	Identify:

	• Human ownership or use of implicated devices, • Context-specific harms, • Unsuspecting data subject expectation and perception should be leading: ◦ Researchers to seek out rules, etiquette, guidelines that govern the virtual space or community, ◦ <u>Researchers should assess whether they are attributing actions to their data subjects.</u> • Self-serving arguments: ◦ Will the perceived benefits indeed materialise, or are they hypothetical? ◦ Research does not necessarily deliver the benefits suggested by researchers, ◦ <u>Relying on legal assumptions (plausible deniability, fair trial) for data subjects.</u> ◦ <u>Legal protections may not exist for data subjects.</u> • Moral harms stemming from vulnerability and ability to object, ◦ Illegitimate treatment even if data subjects are not exposed to any risk, ◦ Groups are treated illegitimately when they are used for research because of their weak position or inability to protest. Understand: • Researcher may not be able to understand or generalise the data subjects' needs, interests, and the context-specific risks of harm, ◦ <u>New ethical issues arise</u> ■ <u>Harms may result from active third parties interpreting initiated information flows from devices and acting upon it</u> ◦ <u>Attributed actions may be incriminating, possibly severely harmful.</u> ◦ <u>Legal compliance may still yield inappropriate and harmful outcomes</u> • Harms may only become apparent after the data collection or publication of the research, • Interests of colleagues may be harmed if (moral) harms occur, • Publication of research may incentivise others researchers to use data from illicit sources, creating potential for further harms, <u>If not for unsuspecting Internet users, the researchers would use volunteers who put themselves in immediate danger.</u> Interpret: • Which standard is the researcher imposing on all data subjects? • Interpret minimal risk standard on a continuum, instead of the static and universal 'harm [...]' no greater than those ordinarily encountered in daily life' standard, ◦ Differentiate between: ■ Direct harm, ■ Indirect harm, ■ Potential harm. • Interpret harm beyond monetary loss and physical injury. • <u>Foreseeable harm:</u> ◦ <u>Requirement to gain insight into types of harms.</u> ■ <u>e.g. Survey research to assess user expectations and perceptions.</u> • <u>Extrinsic/downstream harms:</u> ◦ <u>Counterfactual dependence: would outcome have come to bear without the actions of the researcher?</u> ◦ <u>Substantial cause requirement: Assess whether the degree of causation of negligent act was substantial enough to justify attributing responsibility to the researcher.</u> ◦ <u>Sine qua non: would harm have occurred without the researcher's action</u> ◦ <u>Professional negligence:</u> ■ <u>Harm due to negligence by professional with higher expected knowledge and skill?</u> ■ <u>Should researchers hold themselves to a higher standard than other technical actors in society?</u> ◦ <u>Special duty of care: Assess the impact of actions on implicated third parties.</u>
Emotions as moral tools	Positive: • Benevolent intentions, ◦ Higher-order goal of improving security, ◦ <u>Higher-order goal of enabling the free flow of information.</u> ◦ <u>Higher-order goal of generating data on the functioning of the Internet.</u> Negative: • Possibility of conducting deception research without protocol, • Exploitation of vulnerable persons, ◦ who are not aware of the research, ◦ who cannot exercise their rights, ◦ who are unable to object or protest. • Selfishness: Scavenging on the fruits of crime, ◦ For personal and professional benefit, ◦ Risking reputational loss to discipline.

	• <u>Carelessness: Imposing assumptions of legal defenses.</u> ◦ <u>Plausible deniability. due process. fair trial.</u>
--	--

Autonomy	
Relational approach to morality	Data subjects: • Unknown or misleading identities, ◦ Researchers, ◦ Data subjects. • Status of relationship is unclear if identities are unknown, • Possibility to communicate establishes potential relationship, ◦ Researchers should explore possible channels of communication with data subjects, ◦ Presence of email addresses in dataset allows communication. • <u>Diminished autonomy: Internet users have little knowledge of how much data their devices generate on behalf of third parties.</u> Society: • Researchers can enrol Internet users seemingly at random: ◦ <u>Proxy consent: Researchers ask website or platform owner for consent on their users' behalf.</u> ■ <u>Reciprocity (quid pro quo): Researchers may provide a service to websites and platform owners in return for giving access to the devices of their users.</u> ◦ <u>Capacity to interfere (domination).</u> ◦ Diminishing the ability of all Internet users to choose to take part in research, ■ <u>Worsening choice situation.</u> ◦ <u>Using Internet users as a means to an end.</u> • Lacking clarity whether informed consent establishes autonomy or a relationship meaningfully in research at all, • People's interactions with those who have consented, can be implicated, • Ability by researchers to alter psychological state of Internet users, • Ability by researchers to change Internet users' environments, ◦ <u>Making use of the affordances of the Internet.</u> Research profession: N/A
Contextually-oriented deliberation	Research context: • Can researchers infer the intentions and expectations of users to provide existing data? ◦ Higher-order reason why users provided this data dictates its context and reason for researcher to process further, ◦ Dehumanisation: Responsibility to assess the sociopolitical and human impact on implicated data subjects. • The researcher may be required to make an effort to make themselves known to data subjects, • Difficulty or impossibility to identify vulnerable persons or groups within the research population, • <u>Informed consent requires apprising users about nuanced technical concepts and contextual harms.</u> ◦ Communication about foreseeable risks is not feasible without a contextual analysis, • Notification of rights (e.g. to withdraw) or debrief after deception may not be possible. • If informed consent is not feasible, ◦ Consider alternatives (e.g. opt-out), ◦ <u>Special obligation: Deliberate on the ethical dilemmas, tensions and issues that are specific to the nature and context of the research they are conducting.</u> ◦ Weigh moral gravity of potential rights violation against the social value of research outcomes. • Other applicable laws may prescribe necessity for informed consent. Data subject context: • Difficult or impossible to understand data subject context, • Researcher has responsibility to assess the Intention of data subject: ◦ E.g. to keep data secret. • Responsibility to assess whether the reuse of data exposes data subjects to new risks (indirect) harms. ◦ <u>E.g. active third-parties who may also collected generated data and falsely attribute actions to implicated Internet users.</u>
Responsiveness	Identify: • Starting point: gain informed consent ◦ Are data subjects in a position to give their informed consent? ◦ Uncertainty whether research subjects understand what they consent to, ◦ Competence to provide informed consent may vary,

	○ Vulnerability status (decisional capacity) of research subjects. ○ Legal requirements for informed consent, ○ Can the data processing and purpose be explained to data subjects in layman's terms? ○ Assess burden of consent, ○ Consider alternatives, e.g. proxy person, opt-out. ● Alternative: waive informed consent ○ Overly complex if data subjects are dispersed around the world, ○ <u>Informed consent is not required when it is prohibitive or impractical and there is little appreciable risk of harm.</u> ○ Consider: ■ High costs of gathering informed consent (time & resources), ■ Moral gravity of the potential violation of data subjects' rights, ■ Threat to undermine validity of research. Understand: ● Impossibility to make a judgement of needs if the identity of data subjects is unknown, ○ Obligation to understand the expectations of data usage by data subjects, ● Assess potential harms beyond legal infringements, monetary loss, and physical harm, ● Assess vulnerability of groups in research population, ○ <u>Making use of affordances of the Internet in previously unknown ways may set a precedent and enable further malicious activity as new attack vector is discovered.</u> ● Can the data subjects protest or withdraw from the data collection? Interpret: ● Reasonable effort to obtain informed consent, ○ Assess technical options with which information can be provided to an Internet user and consent can be gained. ● Explore information provision in layman's terms, ● Can informed consent alone safeguard the autonomy of vulnerable groups? ● Consent via proxy: is the person giving consent (e.g. from website or platform owner on behalf of their users) legitimised to do so?
Emotions as moral tools	Positive: ● Respecting autonomy through engagement with data subjects, Negative: ● Outside interaction with a person's device may be a violation of rights, ○ Violation of rights can (also) be considered as a moral harm, ○ Potential violation of the freedom from being subjected to research, ● Violation of autonomy to use data without consent, ○ Violations of dignity, ○ <u>Violations of welfare</u> ● Recklessness: ○ Potential to create further harms beyond the control of researchers, ● Exploitation: ○ No knowledge of research by data subject, ○ Data subjects are used as a means for research because of their weak position or inability to protest. ● <u>Domination:</u> ○ <u>Capacity to interfere arbitrarily.</u> ○ <u>Limiting unsuspecting Internet users choice architecture.</u>

Privacy	
Relational approach to morality	Data subjects: - Relationship to researcher established through networked technology connectivity, - The identity of Internet users is intertwined with their devices due to increasing sociotechnical convergence, - Re-identification in reused datasets may establish a relationship to previously anonymous data subjects, - Data points may be inherently identifiable, - Data points may enable communication (e.g. email addresses), - Data points may identify unique devices or identifiable connectivity sessions (e.g. IP-addresses). - Those with technical capability to collect data, establish the relationship. Society: - Researchers can in theory collect research data about any Internet user, - Without the Internet users need for knowledge or consent for data collection. - Data sources may not be intended for public use, - Even though they are technically accessible, freely or for a fee, - Passive data collection: purpose for sharing data should be leading, - The power to collect data establishes a power hierarchy, - A form of surveillance. - Violations of privacy can be justified by higher-order purposes, - E.g. improving security mechanisms for data on the Internet, - Justifications must be robust, - Negligible risk to data subjects. Research profession: - Data from illicit sources are extensively used in technical Internet research.
Contextually-oriented deliberation	Research context: - Distinction between public and private domains are blurred. - Obligation to reason assess data availability beyond mere legal compliance (e.g. data protection, information privacy, research ethics), - Intended accessibility of data indicates perception of privacy by data subjects, - Higher-order goals and purpose of data usage to infer expectations and perceptions. Data subject context: - For passive data collection, the reuse of data in research context may not be perceived as appropriate. - Consider intention by data subject to keep data secret - Data unlikely intended for research purposes or public exposure - Data can come back to haunt data subjects.
Responsiveness	Identify: - User expectations and perceptions of privacy in their online context, - Respect high-order goals of privacy as part of technology functions, - For passive data collection, identify the purpose for which data is originally shared, - Potential for re-identification after anonymization of dataset. - For active data collection, identify third parties who may also collect generated data, Understand: - Sensitivity of data for data subjects, - Obligation to exercise special caution, - When data subjects do not know they are participating or data is collected about them, <u>When researchers are attributing actions to Internet users without their knowledge, consent, or participation.</u> - No authoritative guidance exists for data use from illicit sources. - Lines between data collection, illicit intrusion into devices, and deception, - For passive data collection from an illicit source, the responsibility for researchers may be unclear. Interpret: - Special caution when collecting data on unsuspecting data subjects, - Technical ability and capacity to collect data does not confer the right to use this data, <u>Technical ability and capacity to generate data flows from unsuspecting internet users' devices may enable third parties to collect identifiable information about users.</u> - Context-sensitive and responsive approach requires special caution.

	• Tension: Data utility vs. purpose limitation, ◦ Do advances for science justify speculative ends of research? • Determine level of privacy to be respected based on: ◦ Group accessibility, ◦ Perceived privacy of data subjects. ◦ User expectations and perceptions of privacy dictate the role a researcher may play in an online space. • Offline analogies should not replace assessment of user expectations.
Emotions as moral tools	Positive: • Control of data, ◦ To maintain dignity. Negative: • Potential for surveillance, • Loss of privacy, ◦ Reported feeling of being dissected like laboratory animals, ◦ Internet users become a means to an end for researchers. • Feeling of being violated.

(6) Analysis

This final chapter ties together the developing conceptual framework and completes the thought-experiment on the intersection of law, ethics and technology. To this end, the chapter is divided into three parts. The first part of the chapter is a literature-based scrutiny of the developing conceptual framework. It answers the third research question: *How might these concepts be reinterpreted for the purposes of research ethics evaluation of technical internet research when considered through the appropriate ethical framework?* This part takes as its starting point the tables of the conceptual framework as developed throughout the case study chapters to assess the original formulations of the Common Rule through a new lens. It seeks clarifications in the ethics of care literature and proposes a new formulation of the challenged legal provisions. The second part provides the practical contribution for the various groups identified in the introduction chapter by answering the main research question: *How ought technical Internet research with unsuspecting data subjects be assessed from the perspective of research ethics?* These groups are (1) research ethics gatekeepers at academic institutions (e.g. IRBs), (2) technical Internet researchers, (3) policymakers or other persons responsible for devising research ethics frameworks for technical Internet research, and (4) the scholarly ethics of care community. The third and final part suggests some fruitful avenues for further research for the re-interpretation of laws when the sociotechnical environments they regulate, have changed significantly.

Literature-based scrutiny

The literature-based scrutiny of this thesis presents challenges of the conceptual framework, as developed in the previous chapters, through the lens of the ethics of

care literature. In addition, it articulates new, re-interpreted legal provisions that are based on this ethical framework. This part is the analytical culmination of the methodology chosen for this thesis. It is also the most complicated part, because an analysis based on decades of theoretical and philosophical literature could be an endless endeavour. Therefore, some limiting decisions are required for the analysis to be useful. So, before the analysis is presented, an overview of the limited theoretical lens and methodology are described below.

This section sketches the theoretical background for law that is based on the ethics of care, rather than the typical justice-based underpinnings. It thereby builds on the description of law and ethics in the second chapter. Further, it describes the environment created by a law based on the ethics of care and the means it incentivises to achieve its end. Responsiveness is identified as the core value that should be promoted in an ethics of care-approach to law, though it is its antonym—unresponsiveness—which requires regulation. Finally, the methodology as described in general terms in the introductory chapter is geared towards the analysis of the concepts and articulation of new legal provisions.

These sections pave the way for the main analysis of this part, which challenges and reinterprets the concepts that have been identified as leading to inappropriate outcomes of the current research ethics review procedure for technical Internet research. The analysis first examines the human subjects concept for technical Internet research. Then, the section on ‘autonomy’ has been expanded to include ‘vulnerability’ and ‘relationships’ and is moved forward, ahead of the sections on harm. Finally, the section on the ‘minimal risk of harm’ and ‘privacy’ have been

combined. These changes reflect the importance of vulnerability and relationships in the conceptual framework, and its overlaps between harm and privacy.

Positioning ethics of care-based law-making

The creation and design of new law requires some theoretical background as part of its methodology. This section describes the aims of a law based on the ethics of care, compares it to the aims of law based on other, more traditional ethical frameworks, and describes the means to achieve the aim. By positioning the theoretical framework for the new provisions in this way, this section builds the foundation for the analysis in the following sections.

The aim of an ethics of care-based law has been articulated by various scholars, and is reflected in the descriptions of care. For example, Bubeck describes care as an activity that meets the needs of others (Bubeck, 1995). Approaches based on care therefore promote values such as sensitivity, trust, and mutual concern (Held, 2005). Conversely, moral harms in the ethics of care literature are described as failures to respond to needs and taking appropriate responsibility (Kittay, 2011).

Jonathan Herring describes how law can be "used powerfully to encourage people to act in a way that will promote caring relationships" (Herring, 2017, p. 164). He argues that a law built on an ethics of care foundation supports an environment within which positive relationships flourish, because responsive attitudes and actions by individuals are incentivised. As an example, he describes that it may not be possible to formulate enforceable laws that prescribe that a nurse must act in a caring way towards patients. However, law can create the "environment and processes that might facilitate compassionate caring relations", just as "contract law

plays a role in creating an environment in which contracts are respected." The theory therefore situates law to incentivise agents to fulfill the responsibilities, duties, and obligations that arise from their "[...] accidental or purposeful embeddedness in particular social contexts, the power dynamics within these situations, and the vulnerability of those for whom the moral agent has a responsibility to care" (Held, 2005).

Similarly, Sevenhuijsen suggests that integrating the values promoted by the ethics of care into an institution or social role (which could be that of the technical Internet researcher), it "will be enriched and thus better able to cope with diversity and plurality, and care will be 'de-romanticized' enabling us to consider its values as political virtues" (Sevenhuijsen, 2003, p. 15). A research ethics law based on the ethics of care should therefore not only incentivise individual researchers to be more caring and responsive about the effect of their work on research subjects, but should result in the institutions of technical Internet research (e.g. computer science departments, conference, corporate R&D departments) to also develop a caring and responsive attitude to their work.

The aims of a care-based legal system for research ethics also can be illustrated by contrasting it to the aims and reasoning methods in existing laws that are typically underpinned by abstract and universalist theories of justice, such as consequentialism and deontology. Justice-based approaches to law primarily promote values such as equality, freedom, and seeks a fair solution between competing individual interest and rights. A research ethics law thereby becomes a tool that enables researchers and human subjects to exercise their rights with regards to each other. This assumes that participants in research are autonomous and can

choose to engage in research studies, and reject experiments that do not serve their interests or needs.

In these research ethics frameworks, researchers are incentivised to reason about their responsibilities in the form of moral accounting (consequentialism) or their compliance with rules (deontology). They need to only ask “is it my right to do X?” and not necessarily to think further, as long as their actions are in compliance with the system and concepts of the law. Such an approach has created the illusion for some researchers that they are acting 'ethically', while my assessments in the case study chapters show that compliance with justice-based legal provisions can yield inappropriate (or 'unethical') outcomes. The confusion that results from legal compliance in technical Internet research has led to a conceptual vacuum (or a policy vacuum (see Moor, 1985)), where it is unclear for a technical Internet researcher how to act responsibly.

A care-approach to law does not necessarily exclude justice-based approaches, but can help to interpret responsibilities in concrete contexts. Held suggests that that ethics of care may help clarify how caring activities should be incorporated when thinking about rights, instead of leaving the law to dictate one's duty (Held, 2005). Kittay further criticises legalistic approaches to morality, when she states that “practices of justice can be inadequate to cover the contextual and narrative complexities of the situation” (Kittay, 2011, p. 53), such as demonstrated with the inappropriate outcomes of the case studies in this thesis. Therefore, Held argues that justice-approaches to moral reasoning should not be a guiding model for moral life. She suggests that an assessment of caring relations can be used to gain deeper understanding of the context within which justice is sought.

Designing ethics of care-based law-making

As described by Herring, the law may not be able to prescribe the precise end it aims to achieve. The law can, however, incentivise the means to achieve a particular environment where specific values are promoted. Herring states that it may not be feasible to impose care in relationships, though caring attitudes can be incentivised (Herring, 2017). Therefore, a research ethics law based on the ethics of care should communicate incentives for researchers and their research ethics review boards to adopt caring attitudes towards relatively vulnerable unsuspecting Internet users who happen to be implicated in technical Internet research. Held proposes that laws based on the ethics of care should shift the researchers' primary question from understanding their rights, to "[w]hat is my proper obligation within the context of this relationship?" (Held, 2005, p. 15). Such a question is unlikely to be answered satisfactorily by researchers individually, however. Research ethics review committees will have a guiding role to play to ensure caring incentives are upheld.

Slote suggests considering incentives to care not as prescriptive, either, but to penalise *uncaringness* (Slote, 2007). In the context of this thesis, the idea is that care by researchers can and should be expected by research subjects, and uncaring attitudes that lead to moral harms must be regulated. Brannelly similarly states that it is the unwillingness of researchers to engage outside of a research protocol based on the ethics of care, which is likely to violate the trust of research subjects (Brannelly, 2018). The research themes of the thesis indicate that a responsive attitude is key to a caring attitude. Responsiveness requires the researcher to identify the needs of the other in their context, to understand the identified needs, and to interpret them into the research context. Therefore, to create a caring environment, researchers need to

be incentivised to be responsive to the needs of their data subjects. In short, responsiveness can be promoted by penalising *unresponsive* attitudes. Demonstrating responsiveness would thereby become the threshold to pass research ethics review, based on an ethics of care framework.

Responsiveness as threshold

The third research theme in chapter 2 constructed a practical approach to how scholars have conceptualised responsiveness. As part of the literature-based scrutiny, this section describes the theoretical background to responsiveness as a value, and how it may be operationalised specifically in the context of research ethics review. It builds on the formulation of the third research theme, which describes responsiveness as:

1. The responsibility to identify the needs of vulnerable users who are implicated in a context and thus have a relationship with the researcher,
2. The effort to understand these users' needs and interests, and
3. The interpretation and accommodation of their needs in the (technical) research design.

These steps reflect an understanding of responsiveness as an effort by the more powerful actor in a relationship to gain knowledge of the social and political, as well as a personal contexts of the more vulnerable actors, and to understand their point of view (Noddings, 1982; Tronto, 1993). Responsiveness requires not only good intentions, but to actively gain knowledge of the research subjects' contexts, an assessment of their needs, and thoughtful judgements in the research design (Tronto, 1993; Sevenhuijsen, 2003). This moral obligation stems from the power asymmetry in

the research relationship, and the resulting dependence and vulnerability of the research subject (Nortvedt, Hem, & Skirbekk, 2011). Failure to act responsibly towards the vulnerable and dependent, results in moral harm (Kittay, 2011).

Scholars suggest that an assessment of relevant needs requires a deliberative space, which connects the morality of the research context to that of its participants (Tronto, 1993; Brannelly, 2018). Visse et. al describe a new role for an ethics review committee as a 'responsive evaluator', who facilitates approaches to elicit, articulate and connect the voices of people with policy, such as a research ethics assessment (Visse, Abma, & Widdershoven, 2015). This approach is discussed in more detail in the next part of this chapter. However, for this section it is suffice to present the responsive evaluator to enable researchers to view the morality of their work from the point of view of their research subjects (Tronto, 1993) and to jointly respond to the newly unearthed knowledge (Stilgoe, Owen, & Macnaghten, 2013). In conclusion, to design an ethics of care-based research ethics review law for technical Internet research, a research ethics review committee should actively promote and facilitate the responsiveness of their researchers towards their implicated data subjects, with the threat of withholding approval for research if the researchers are unresponsive.

Methodology to articulate ethics of care-based research ethics law

The introductory chapter described the structured case methodology by Carroll and Swatman, which has been used throughout this thesis. The final chapter of their methodology is literature-based scrutiny. They refer to this stage of their methodology as "enfolding literature", which is the final stage of the "spiral towards understanding" (see the first chapter for a more detailed description of the full thesis methodology). It involves reflection on the developed conceptual framework,

comparisons to literature of the underlying theory, and critical assessments of the findings so far (Carroll & Swatman, 2000).

Carroll and Swatman designed their methodology to expand existing theory, or to create new theories, specifically in qualitative information system's research. The original structured case methodology uses this stage in the research process to articulate extensions or strengthening of the existing theory. This chapter adapts this part of the methodology from a theory-building tool to a means to re-interpret legal provisions through a new theoretical underpinning. The following steps are applied in this section:

- First, the Common Rule provisions that are challenged throughout the development of the conceptual framework are re-stated and dissected, with the aim to uncover their underlying values and aims.
- Second, the identified challenges to the existing provisions are articulated in terms of the ethics of care literature and the research themes of this thesis. The tables of the conceptual framework are the starting point for analysis.
- Third, at the end of each sub-section, a new legal provision is designed that responds to the challenges and incorporates the analysis of the first and second step, outlined above.

The newly articulated provisions intentionally leave room for interpretation by researchers and their research ethics committees. Some concepts are not defined precisely on purpose, leaving discretion to the ethics committee to interpret the appropriate level of responsiveness. This is to avoid a compliance framework, such as the current Common Rule. Technical Internet research will continue to be faced

with (unforeseen) moral dilemmas, which will be the result of experimenting with complex technologies that impact the pluralistic nature of human societies on a global scale. The new provisions offer IRBs starting points to request or gather more information in order to determine whether the researchers do, indeed, demonstrate a responsive attitude.

Human Subjects

For the human subject model of research ethics, a meaningful definition of the concept of *human subjects* is of central importance. After all, if a data point represents a human subject, the Common Rule is triggered, and its provisions applied to the proposed study by an IRB. However, the development of the conceptual framework throughout this thesis demonstrates that it is not always clear when a data point that is collected for the purpose of technical Internet research represents, in fact, a human. This section re-visits the Common Rule definition of human subjects, traces some of the challenges to the concept throughout the development of the conceptual framework and proposes a revised articulation of the provision.

Common Rule

Human subjects are defined in §46.102(f) of the Common Rule as follows:

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

This provision defines the concept of human subject and establishes a research relationship between researcher (or ‘investigator’) and human subject. The following assumptions can be deduced from this provision, through the lens of the ethics of care terminology:

Particular or generalised other

The Common rule takes as a starting point a research relationship between a particular researcher and particular human subject. The formulation in §46.102(f) specifies ‘a living individual’ and their ‘identifiable private information’, which assumes that the particular human subject is known (or knowable) to the researcher. The Common Rule does not contain provisions about research conducted on a generalised group, such as a random and unidentified group of Internet users that have specific accounts or devices.

Proximate and distant other

The Common Rule further differentiates between proximate and distant data human subjects. Human subjects are considered to be proximate to researchers when data is collected “through an intervention or interaction with the individual”.

Interpretations provided in the Common Rule are explicitly directed towards the clinical collection of medical data, both physically (e.g. venipuncture) and through manipulating a human subject's environment. A spatial distance between researcher and data subject would complicate the assumed data collection procedures.

Data subjects may also be distant when the researcher obtains ‘identifiable private information’. The interpretation in the Common Rule specifies that this may include data about ‘behavior that occurs in a context in which an individual can reasonably

expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example, a medical record)'.

Therefore, the inclusion of distant data subjects appears to be limited to the reuse of other, possibly sensitive data.

Challenges from conceptual framework

The conceptual framework challenges the definition of human subjects in §46.102(f) and its assumptions in two fundamental ways. These are (1) establishing research relationships to distant and generalised research subjects and (2) the humanisation of technical data points. This section describes these two challenges and provides the analyses through which the provision may be interpreted specifically for technical Internet research.

(1) Establishing research relationships to distant and generalised research subjects

The Common Rule assumes a research relationship to a proximate human subject, whereby an 'intervention or interaction' is performed in person and within a physical research setting, or by the reuse of existing data. This assumption may hold true for traditional research in the medical and behavioural sciences, though is no longer a useful standard for the technically mediated research relationships between technical Internet researchers and their data subjects. The case study chapters of this thesis describe research relationships that are established through technical means, at a distance, and with undefined research populations.

Researchers can make use of the affordances of the Internet to collect data from virtually any (connected) location, so spatial distance does not matter. Most, if not all

human subjects implicated in technical Internet research may therefore be characterised as 'distant' others, as opposed to the proximate others assumed by the Common Rule. The conceptual framework and literature on the research ethics of social science Internet research identify the 'distance principle' for research with distant human subjects. Aycock and Buchanan describe the phenomenon as when the distance between researchers and data subjects increases, researchers are less likely to consider their data points as representing humans, and therefore are less likely to consider their research as involving human subjects (Aycock, Buchanan, Dexter, & Dittrich, 2011; Buchanan et al., 2011). Further, Slote observes that "[s]patial distance and (decreasing) empathy do in fact correlate with one another across a wide range of cases, and that very fact may have helped to obscure the role empathy potentially has in explaining the sorts of distinctions people intuitively, or commonsensically, want to make [...]" (Slote, 2007, p. 21-22).

Further, research relationships in technical Internet research are established as a result of people owning or operating devices that are connected to the same computer network (i.e. the Internet). However, the data collection mechanisms do not purposefully identify or differentiate between Internet users in the research population. Instead, human subjects are implicated in research studies somewhat by coincidence, through password dumps (existing data) or by manipulating the devices owned by Internet users who visit a website that happened to include an invisible data collection mechanism (intervention). Researchers target a more generalised group, without regard for the individual characteristics. Therefore, the research relationship is not between proximate and particular persons, as assumed by the Common Rule, but between distant and generalised groups. A revised

standard is required to incorporate this distinction to include implicated these human subjects in research ethics review.

(2) Humanisation of distant and general data points

The devices and online accounts implicated in these studies are intimately linked to the identities of their human operators or owners. The data collected may reveal personal behaviour and preferences, such as websites visited, or accounts operated. The data points collected for the technical Internet research are therefore not merely technical. When researchers miss, disregard, or forget the humans represented by data points in their research dataset, the data points are effectively dehumanised. Therefore, what may be considered as merely technical data points by Internet researchers (e.g. an IP-address or an email address), must be scrutinised further to uncover whether it involves data about humans, too. Further, some of the data collection mechanisms allow another person (e.g. the researcher) to take control of the implicated Internet users' accounts or devices and attribute actions to them, be it directly or indirectly.

A responsive approach to the human subject concept would therefore begin by adopt the attitudes described by Nel Noddings, of being responsive to the needs of those they care for, and being responsive to their needs (Noddings, 1982). The moral responsibility for implicated data subjects is derived what Held describes as the accidental or purposeful embeddedness in particular social contexts, the power dynamics within these situations, and the vulnerability of those for whom the moral agent has a responsibility to care (Held, 2005). Researchers should therefore identify whether their technical data points are effectively tied to humans. If so, the researchers should demonstrate an understanding of what their data collection

mechanism and overall research procedure will mean for these implicated humans. Tula Brannelly describes an ethics of care approach to research as connecting with the morality of those affected by the issues under investigation (Brannelly, 2018). Further, Magda Lewis argues that the failure to critically assess a research's "social/political engagements in a human context" should be a key consideration in refusing to let a proposed study pass research ethics review (Lewis, 2008, p. 693). This relates to the idea of responsiveness being the critical threshold for passing a research ethics review procedure under an ethics of care regime. Therefore, the responsive approach requires a general understanding by the researcher of who the data subjects are, their context, and the informational sensitivities they experience as it relates to the collection of research data.

Further, researchers should actively consider adapting their data collection and research procedures, based on the knowledge about the human subjects and their environments. Dominique Rivière suggests an IRB can support researchers to interpret these ethical challenges by humanising their data subjects. For example, the IRB should ask the researchers to make educated guesses about the social and political engagements of their human subjects, and how the likely impacts and arising tensions might be handled (Rivière, 2011). While it may not be possible to identify each data subject individually and determine their needs specifically, experience from approaches in welfare or security show that the needs can be determined on a more general level (Engster, 2015; Robinson, 2011).

In summary, a responsive approach to distant and generalised human subjects requires researchers to first identify who is likely implicated. Then, researchers need to demonstrate an effort to understand the impact of the research on their data

subjects in their own, local contexts. Finally, researchers should demonstrate how the new knowledge is interpreted in the research context, and justify how the research procedure is designed to avoid being dominating, oppressive, and exploitative.

Articulation of a new provision

The two challenges to the current formulation of §46.102(f) of the Common Rule create space for a reinterpretation of this provision. This subsection provides an answer to research question 3 (“How might these concepts be reinterpreted for the purposes of research ethics evaluation of technical internet research when considered through the appropriate ethical framework?”) for the concept of human subjects. The new provision uncovers unresponsive attitudes by researchers in a three-step process.

First, the first paragraph identifies a research relationship through technical means explicitly, beyond the existing approaches of the Common Rule. The second paragraph requires researchers understand the characteristics of their human subjects. If the characteristics are unknown, their data subjects are automatically categorised as 'vulnerable'. This imposes stricter requirements for researchers, which are discussed in subsequent sections. Finally, the third paragraph offers Research Ethics Committees (RECs, such as IRBs) the power to determine whether data points in a research study represent human subjects, and to what extent they should be considered as a vulnerable population. The concept of 'vulnerability' links this human subject provision to the other new provisions that are designed and discussed in further sections of this chapter. The italicised words are clarified underneath the newly proposed provision.

§A - Research relationship with human subjects

(1) A human subject is a human person with whom an investigator conducting research establishes a *research relationship*:

- By collecting data through personal *interaction* with the human subject, either in person or mediated through technical means,
- By collecting data through *interactions* or *interventions* with a *device* owned or operated by the person,
- By collecting existing data which contains identifiable information about the human subject, either directly or when combined with *reasonably available* auxiliary data.

(2) The REC may demand investigators demonstrate, to the extent *reasonably possible*, an understanding of the human subjects and their environments. This includes:

- An appropriate generalisation of the foreseen human subject's *relevant characteristics*,
- Relevant informational sensitivities in their *sociopolitical environment*,

If the characteristics and sensitivities cannot be known in advance, the IRB human subjects must be considered as a *vulnerable population*, unless the investigators can justify why this characterisation is inappropriate.

(3) The REC determines whether a research study involves human subjects, and to what extent they should be categorised as a vulnerable population. To this end, the REC has the discretion to require investigators to provide additional information

and justification in writing or in person, for which §B and § C provide further guidance and requirements. If the IRB determines that a research study does not include human subjects, the research may be described as *mere technical research*.

Research relationship: The relationship between the researcher (or investigator) and the individual human subjects in their research database.

Interaction: An interaction with a person or device implies that some form of communication is taking place.

Intervention: An intervention with a device implies that the researcher manipulates the functionality of the device in some form.

Device: A device is an information technology, which processes information and may be connected to a network.

REC: A Research Ethics Committee, such as an IRB.

Reasonably available: Data is reasonably available when the researcher, a potential attacker, or a potentially interested third-party who may take note of the data collection mechanism, can use the research data to identify individual data subjects.

Reasonably possible: The sociopolitical assessment should not be an absolute, though required on a best-effort basis. The REC determines whether they are satisfied with the extent to which human subjects have been described and may support the process.

Relevant characteristics: The relevant characteristics include, but are not limited to, the social demographics of the human subjects, their geographic location, their expected level of understanding of the data collection process (if informed).

Social environment: The human subjects' sociopolitical environment is composed of many variables, including but not limited to government institutions, religious

institutions, the independence of the judiciary, emerging social movements, and various other sources of social power.

Vulnerable population: Vulnerable populations are those with diminished or non-existing autonomy. The concept is clarified further in §B.

Mere technical research: When a research proposal studies incorporates only data points are not directly or indirectly linked to humans, but only to technical artefacts, the study may be considered as a 'mere technical study'.

Autonomy and vulnerability

Autonomy, understood as self-determination, is a core value of the justice-based approaches to morality. The Common Rule operationalises research subjects' autonomy through a required informed consent procedure. A relational approach to morality, as proposed by the ethics of care, does not replace or shed the value of autonomy. However, the level of self-determination of more vulnerable parties informs responsibilities, obligations, and duties that arise from hierarchy and power balances.

For the purpose of the literature-based scrutiny, the concepts of vulnerability and dependency have been added to this section. This reflects how autonomy is balanced with vulnerability and dependency. The degree of autonomy, vulnerability, and dependency that exist within a relationship, inform the level of caring attitudes and actions required by researchers. Further, this section has been moved forward, ahead of the section on harm and privacy, because of the central importance of autonomy, vulnerability, and dependency in an ethics of care assessment.

This section presents how autonomy and vulnerability and articulated and regulated in the Common Rule. Dependency is not mentioned in this law. Then, two distinct challenges to the Common Rule are discussed. First, the research relationship towards unsuspecting and implicated human subjects is characterised in terms of the ethics of care, including their *relational* autonomy, (inherent) vulnerabilities and dependencies, and the negative research relationship characterisations in technical Internet research. Second, the primacy of informed consent as a mechanism to establish a balanced research relationship is challenged. This is, in part, because the focus of this thesis on technical Internet research that collects data from unsuspecting Internet users, who did not consent and have no knowledge of the research. These challenges reposition the mechanism for morality in research firmly back onto the researcher, incentivizing them to align their interests with those of their participants, society, and the research profession.

The proposed provision reflects these challenges and incentivises researchers and RECs to collaborate in adopting a responsive attitude. The provision requires researchers to justify how their relational approach yields a positive relational outcome, where interests of researchers and researched are aligned. It further requires researchers to explore mechanisms to inform human subjects as a means to establish an explicit relationship, and to take conscious steps to protect potentially vulnerable populations in their research. These steps contribute to a responsive attitude, along with the other provisions proposed in this chapter.

Common Rule

The Common Rule's starting point to autonomy is the assumption—based on its underlying universalist and abstract ethical theories—that participants in a research

study ought to consent to the research relationship to maintain their self-determination. An informed consent procedure is therefore taken as the mechanism to establish autonomy. When people are not in a position to consent to participating in a research study, for example when they have diminished autonomy or are considered to be vulnerable in some other way, the Common Rule imposes some precautions from researchers. This section first briefly presents the informed consent requirement as the mechanism by which autonomy is supposed to be achieved in research ethics. Further, this section describes how the Common Rule conceptualises vulnerability, and which requirements it imposes on researchers who include these groups within their research data collections.

Informed consent

The Common Rule describes the scope of the informed consent requirement in three provisions. §46.109(b) states the starting point of the Common Rule, that an IRB must require researchers to provide information to human subjects. Subsection (c) of this provision requires an IRB to collect documentation when it has waived the researchers' requirement to collect informed consent. Further, §46.116 describes the requirements for an informed consent procedure, and restates its importance for research ethics review:

Except as provided elsewhere in this policy, no investigator may involve a human being as a subject in research covered by this policy unless the investigator has obtained the legally effective informed consent of the subject or the subject's legally authorized representative. [...]

An informed consent procedure is typically prescribed in research ethics frameworks. Similar provisions are found in the Nuremberg Code and the Helsinki

Declaration, which are the foundational legal frameworks for the human subject model of research. The source of this requirement is Kant's deontological theory (called the categorical imperative), which research ethics scholars have interpreted to prescribe that "we should not treat people as means without their consent" (Veatch, 1987, p. 56). Similarly, scholars such as Faden and Beauchamp understood autonomy in the realm of research ethics as "restrict[ing] human subject research to those who (1) intentionally, (2) with understanding, and (3) without controlling influences agree to be subjects" (Faden & Beauchamp, 1986, p. 238). The Belmont Report, which initiated the Common Rule, expresses that "respect for persons demands that subjects enter into the research voluntarily and with adequate information" (Commission, 1978, p. 4). The report describes the three elements that constitute the concept of informed consent:

- **Information:** an expectation to disclose the research procedure, its purpose, anticipated benefits and identified risks, and the opportunity to ask questions and withdraw from the research at any time.
- **Comprehension:** giving the data subject the ability to understand the information and make an informed choice, by presenting it in an organised manner, allowing sufficient time for consideration, and allowing questions to be asked.
- **Voluntariness:** consent that is free from coercion (e.g., the threat of harm) and undue influence (e.g., offering an inappropriate or excessive reward) (Commission, 1978).

Finally, §46.116(d) describes some exceptions to the requirement of informed consent:

- The research involves no more than minimal risk to the subjects;
- The waiver or alteration will not adversely affect the rights and welfare of the subjects;
- The research could not practicably be carried out without the waiver or alteration; and,
- Whenever appropriate, the subjects will be provided with additional pertinent information after participation.

Vulnerability

Vulnerability in research ethics is typically assigned to groups with diminished autonomy (Commission, 1978). When a person is not in a position to make judgements and is thus considered to be vulnerable, the researcher has an obligation to protect that person based on a calculation of the benefits and risk of harm. Using the same themes, though in more detail, the ‘European Textbook on Ethics in Research’ identifies as vulnerable those (1) who lack competence to protect their interests in an informed consent procedure, (2) whose voluntariness is compromised, and (3) whose physical or psychological condition leaves them especially liable to harm (Hughes et al., 2010).

The Common Rule provides a non-exhaustive list of categories that legislators consider to be vulnerable in §46.107 and §46.111(a)(3), which is articulated as follows:

“[...] research that involves a vulnerable category of subjects, such as children, prisoners, pregnant women, or handicapped or mentally disabled persons, or economically or educationally disadvantaged persons”

The Common Rule imposes three requirements on researchers who include vulnerable populations in their research:

1. §46.111(a)(3) requires researchers and IRBs to ensure that the experimental set up is “conizant of the special problems of research involving vulnerable populations [...]”.
2. Sub-section (b) of this provision further requires that “[...] additional safeguards have been included in the study to protect the rights and welfare of these subjects.”
3. Finally, §46.107 requires IRBs to consider including representatives from vulnerable groups, or those with explicit experience working with those vulnerable groups, within their committee.

The list of vulnerable groups in §46.107 and §46.111(a)(3) is intended to be non-exhaustive, though the Common Rule provides procedural guidance for the additional requirements when including human subjects from the listed groups. For example, the Subparts B, C, and D of the Common Rule determine which officers, civil servants, or family members must give approval for research on the named groups. The examples in the subparts concern mostly medical and behavioural experimentation, and do not (explicitly) anticipate other types of research.

Challenges from conceptual framework

This section presents two challenges, which shape the new provision. First, this section challenges the centrality of autonomy as the primary source of moral obligation. The ethics of care literature suggests shifting the focus of moral deliberation towards a more relational view of autonomy, which takes into account

the degrees of vulnerability and dependency that result from a (research) relationship. Characterising a research relationship by the level of vulnerability and dependency experienced by actors, informs the duties, responsibilities, and obligations for the more powerful actor in the relationship.

Second, this section challenges the informed consent mechanism to fulfil moral duties. This is especially pertinent in a research environment where the identities of researchers and participants may not be known, and significant knowledge and power imbalances may exist in the relationship. The ethics of care alludes to a participatory design approach, where participants are consulted about their needs and suggestions in the research design. Further, this section suggests that several channels of communication exist between researchers and participants in technical Internet research, and that information provision may be more important than gaining consent.

(1) From individualistic autonomy to relational autonomy

The Common Rule presents an individualistic approach to autonomy, which is underpinned by abstract and universalist ethical theories that promote value such as independence, self-interest, and rational gain-maximising decision-making (Dove et al., 2017). The starting point for most research ethics review procedures is to ensure participants maintain their individual autonomy. Informed consent is the assigned mechanism whereby potential participants may choose to explicitly establish a research relationship. However, imposing this mechanism for moral deliberation on potential participants assumes that participants are knowledgeable of the subject matter or discipline, and are competent ethical reasoners (Haker, 2013). In the realm of technical Internet research, implicated participants are most likely not

knowledgeable of disciplines such as computer science or network engineering, and most likely are not competent ethicists, either. This first challenge argues that not individual autonomy, but a more relational view of autonomy yields morally more appropriate outcomes when conducting technical Internet research.

Different types of research relationships

Research relationships in most academic disciplines are hierarchical in nature, where researchers generally occupy a position of superior knowledge of the subject matter and experimentation methods with regards to their data subjects. Further, researchers create the risks of harm to which participants are exposed. To understand these risks and operationalise their autonomy meaningfully, research participants are dependent on the information provision by researchers in a form they can understand (Chadwick, Ten Have, & Meslin, 2011; Spruit, van de Poel, & Doorn, 2016). In the social, behavioural, and biomedical sciences, research relationships are typically between proximate and particular others. This allows for a personal and dialogical information exchange, which enables research participants to make an assessment of whether they want to be involved in a study. To assure sufficient participation for their studies, research professionals thereby have an incentive for researchers to align their needs and interests with those of the foreseen research participants. The mechanisms to operationalise individual autonomy therefore make sense in the types of research relationships foreseen by the Common Rule.

The accumulated analysis in the conceptual framework suggests that the research relationship between technical Internet researchers and their unsuspecting data subjects differs from the type of relationship envisaged in the Common Rule. As

discussed in the previous section on human subjects, technical Internet researchers are typically at a distance from the data subjects, and the identities (or even their activities) are unknown to participants. The incentive to align needs and interests does not exist for technical Internet researchers in the same, direct way it does for researchers in the social, behavioural, and biomedical sciences. Technical Internet researchers can collect data at a distance and somewhat invisibly, without the need for a dialogue before the data collection can commence. As a result, research participants involved or implicated in technical Internet research are less able to operationalise their autonomy in the research ethics process, if at all.

Relational approaches to autonomy

The failure of individual autonomy to operationalise self-determination for research participants, invites an exploration of other approaches to autonomy. Literature on the ethics of care has described a pluralistic approach to self-determination, called relational autonomy. A systematic literature review of relational autonomy has found that no consensus exists on the concept, and that most articulations of relational autonomy are reactionary, arguing against the abundant assumption of individual autonomy in existing laws (Gómez-Vírseda, de Maeseneer, & Gastmans, 2019). However, some constitutive elements can be discerned.

Relational autonomy emphasises the interdependency of people, as opposed to the assumption of independence in individual autonomy (Haker, 2013; Osuji, 2018). A person's identity, interests, needs, and autonomy are shaped by the social and institutional contexts in which they are embedded (Dove et al., 2017). Therefore, people exercise their autonomy not merely out of self-interest, but also with regard to others. People tend to be less selfish than universalist and abstract ethical theories

suggest (Walter & Ross, 2014). This view of autonomy resists the idea that society is constituted of individuals who are exercising their autonomy independent of each other, as though social ties do not exist (Held, 2006).

In a relational approach to autonomy, responsiveness to (distant) others therefore requires an understanding of both their social environment, and their resulting vulnerability and dependence within the research context. Most literature about relational autonomy is written in the field of bioethics, where authors consider the personal relations of patients to be guiding for their autonomy (e.g. with family, friends, colleagues). Gómez-Vírseda et al. note that autonomy is considered to be a value that is best understood in degrees and along a dynamic continuum, where power imbalances and inequities are determining factors (Gómez-Vírseda et al., 2019).

Vulnerability and dependency determine degrees of relational autonomy

For this thesis, relational autonomy is determined by the degree of vulnerability and dependency of the research participant with regards to the technical Internet researcher. These relate closely to power imbalance and inequity, as suggested by Gómez-Vírseda et al. For example, Spruit et al. clarify that “[t]he existence of power is almost by definition relational: Power exists only by virtue of there being somebody to influence or control with that power” (Spruit et al., 2016, p. 129). To be powerful someone else must be vulnerable. Conversely, Luna and Vanderpoel describe vulnerability as relational, too, stating the “[v]ulnerabilities can only be discovered by examining an individual in context [...]” (Luna, F. & Vanderpoel, 2013, p. 325). Dependency is an inequity and the result of a power imbalance and vulnerability status. Similarly, scholars writing about research ethics typically

classify research participants with diminished autonomy as vulnerable (Bracken-Roche et al., 2017; Capron, 2014; Marco, 2004). The ethics of care literature further describes the vulnerable as dependent within a relationship that is characterised by hierarchy and power imbalances. (Barnes & Brannelly, 2015; Bracken-Roche et al., 2017; Held, 2005).

The degrees of vulnerability, dependency, and relational autonomy within a relationship prescribe the required caring attitudes. Collins proposed a slogan for the ethics of care, stating that "dependency relationships generate responsibilities" (Collins, 2015, p. 2). It is the status degree of type of dependency that characterises a relationship, which in turn prescribes the responsive duties, responsibilities, and obligations that arise from the relationship. Similarly, Haker argues that the vulnerable and dependent person's "right to be cared for (i.e. to be assisted in her autonomy, flourishing or well-being) transforms into a positive duty: it is not enough only to passively respect the other's needs but they must be met by way of action, response. This response, however, must be acted out as 'responsible response'" (Haker, 2013, p. 47). In other words, the positive duty is a duty of care and responsive attitude towards research participants.

The concept of vulnerability is commonly criticised in research ethics literature, however, for being too vague and not offering clear guidance for policymakers, researchers and ethics review boards. In fact, many ethics guidelines and policies employ the concept of vulnerability, but very few define it. Bracken-Roche et al. conclude that the concept of vulnerability is therefore "inconsistent across policies and guidelines" (Bracken-Roche et al., 2017, p. 6). These authors propose a dilemma for the definition of vulnerability, which is especially pertinent for the assessment of

technical Internet research: a definition of vulnerability must be broad enough to capture those in need of additional attention and protection beyond the norm. However, an overly broad definition could capture all research participants. The next paragraphs demonstrate that an ethics of care and relational autonomy approach is useful to determine the degree of vulnerability and dependency, and to determine where to pinpoint research ethics scrutiny.

Vulnerability and dependency in case studies

The Internet users in the case studies of this thesis are vulnerable to the actions of researchers, and dependent on the researcher's attitude towards their datasets and experimental setups. Their research relationships are characterised by a string of negative relational descriptors. For example, the relationships are *dominating*, because the researcher may limit the choice situation of all (implicated) Internet users by collecting their data or interfering with their devices arbitrarily (Pettit, 1997). The research relationships are also *exploitative*, because the researchers use their position of superior knowledge of the affordances of the Internet to use Internet users (or their devices) as a means to their own research ends (Cohen & Lynch, 2014; Kerstein, 2019). For example, the researchers collected data that was intended to be kept secret by their data subjects (ie. passwords), and made use of devices for their own ends.

Further, the relationships are *oppressive*. Oppression itself is elusive as a concept, though can be considered broadly synonymous with 'unjust social arrangements' (Smith, 2015). The second case study, specifically, is exemplary of an unjust social arrangement that is inherent in technical Internet research: Powerful researchers may manipulate the devices of users to collect new data, while potentially

attributing actions to them (ie. requesting censored websites). The consequences for the research relationship are designed to be beneficial for the researchers, while the unsuspecting and uninformed data subjects have little, if anything, to gain. Research participants in both cases were not offered the opportunity to reflect on their participation in the studies and may be harmed in ways unforeseen or unaccounted for by the researchers, especially if the researchers do not take an interest to avoid any negative consequences.

(2) Shift mechanism for morality to researcher

The Common Rule prescribes informed consent as the mechanism to operationalise the central values of research ethics. This places the burden of moral decision-making on potential research participants. When the informed consent requirement is waived, the decision to include certain populations is placed on the IRB and whomever they may consult. Safeguarding the autonomy of research participants is therefore not a primarily a duty that falls on the researcher. This second challenge suggests shifting the burden of moral decision-making more firmly onto the researchers themselves, requiring a responsive attitude to their vulnerable and dependent research subjects. This shift incentivises researchers to align their interests with those of their participants, as intended by the informed consent mechanism in research relationships with more proximate and particular others.

Waiver for consent side-steps central values of Common Rule law

The case studies in this thesis are representative of a larger body of technical Internet research where researchers sought (and gained) a waiver for informed consent. Researchers typically argue that their "research could not practically be carried out without the waiver", which is one of the grounds to waive the requirement in

§46.116(d). Indeed, research populations are not easily identifiable in technical Internet research and there may be no direct contact with participants. Further, even if the research participants were identifiable, researchers claim they would struggle to inform them meaningfully, due to potential differences in language and technical knowledge.

The central aim of the Common Rule is to operationalise autonomy and self-determination by research participants. A waiver for informed consent under the Common Rule procedure, undermines this moral aim of the research ethics law. According to §46.116(d) a waiver may only be granted when the research does not impact the rights and welfare of research subjects, or does not pose more than minimal risk for them. However, the conceptual framework of this thesis presents a larger range of harms as a result of technical Internet research than those typically considered by an IRB. These include downstream harms to participants, moral harms that may lead to a loss of societal trust in the research discipline, and unforeseen injuries to society and the research profession. The new risks of harm and disparate impacts on rights and welfare of implicated research subjects in many different contexts, the waiver for informed consent leaves research participants vulnerable and dependent on the design choices made by researchers. By allowing researchers to perform experiments and collect data without the knowledge and approval of data subjects, IRBs leave implicated participants with a diminished autonomy. Their autonomy is typically not safeguarded in technical Internet research that is conducted based on a waiver for informed consent.

A more valid reason to waive the informed consent requirement in technical Internet research within the system of the Common Rule, would be an argument based on

deception research. As described in the second chapter, the requirements for deception research whereby it is crucial to withhold information about an experiment or study from participants until after the experiment and data collection have been completed (Elms, Erwin, Gendin, & Kleiman, 1994). Research ethics committees are typically more vigilant to the inclusion of vulnerable persons in deception research, or the extent to which research participants become vulnerable as a result of the experimental design. Baumrind's criticism of deception research (also mentioned in chapter 2 of this thesis) serves as a pertinent concern for research ethics review committees assessing technical Internet research proposals. She considers that, in most cases, the incremental knowledge gained from deception research does not outweigh the fundamental obligations placed on researchers to protect the welfare of their research subjects, and to be trustworthy actors in society (Baumrind, 1985; Baumrind, 2015).

Shifting the mechanism towards the researcher

Taking into account Baumrind's criticism, it may be plausibly argued that most technical Internet research does not provide the knowledge gains that justify potential harms and negative research relationships as described in the conceptual framework of this thesis. Focusing on individual consent, or waivers for consent, does not yield appropriate moral outcomes in technical Internet research.

The next paragraphs describe how an ethics of care approach shifts the moral responsibility from potential participants and IRBs to the researchers themselves. Interestingly, the ethics of care literature argues that autonomy is one of many positive outcomes of a relationship but should not necessarily be the moral focus of relationships. A positive research relationship can therefore exist with a mere

information provision by the researcher, using the channels that are available to them. This could be strengthened with an opt-out procedure, similar to medical research. Further, the ethics of care literature promotes dialogue and an understanding and identification of vulnerable person's needs. Such a consultative approach is akin to the existing sub-discipline called participatory design. Both approaches are described and feed into the new provision at the end of this section.

Information provision and opt-out

Gaining informed consent may be meaningless in some situations. When participants do not understand the subject matter, researchers will have successfully complied with the Common Rule requirement to gain consent, though the participant may be none the wiser about the research, data collection, and any risks of harm to them when they consent to being part of a study. Further, given the global nature of the Internet and the potentially hugely diverse research population implicated in a study, an information provision is unlikely to describe the nature consequences to individuals (Johnson & Powers, 2005). In such situations, an informed consent procedure would, in fact, not demonstrate a responsive approach to the needs and interests of research subjects.

Researchers and ethics boards could approach the provision of information and gaining consent differently. The distance between researcher and participant in technical Internet research undercuts the assumption of the Common Rule that informed consent is a dialogical process, where the researcher may tailor their information provision to the level of understanding of the participants.

However, the networked nature of the Internet may limit the effects of distance, as researchers typically do have one or more channels of communication open to them. Researchers in the first case study collected databases containing an email address for each of their research subjects. The researchers could have sent an email to each participant with basic information and links to websites containing more information about the study. Similarly, researchers in the second case study designed their widget to be purposely invisible to the visitors of a particular website. Instead, they could have opted for clear visibility and provide information about the data collection mechanism. Other technical Internet research studies may have different channels of communication open to them, which can be explored.

Further, gaining explicit informed consent can be impractical, even if channels of communication are available. It would be too quick for a research ethics board to conclude that the whole mechanism should be waived. In the two examples of this thesis, the researchers could have opted for an opt-out mechanism. For example, the email could have been sent, or widget presented, displaying a warning and some basic information about the study and present a link to an informative website. Potential participants may then choose an opt-out option, if they do not wish to take part in the study. These strategies are discussed in chapter 4 of this thesis as being employed often in medical science (see for example (Vellinga et al., 2011; Junghans et al., 2005)). They do not necessarily lead to a large drop out rate (Hewison & Haines, 2006). An active role exists for research ethics committees to help researchers identify and understand whether an information provision and an opt-out is sufficient to meet the needs and interests of implicated human subjects on a per case basis (Visse et al., 2015; Spruit et al., 2016).

Consultation and participatory design

By opening the channel of communication, researchers can invite (potentially) implicated human subjects to voice their concerns, express their needs, and possibly even suggest changes to the experimental design. This is a consultative approach that is encouraged in the ethics of care (Po-Wah, 2002) and other discourse-based approaches to ethics (Haker, 2013). Several design procedures have developed over the last decades that incorporate others and include their points of view, which are so far unexplored in technical Internet research. For example, researchers and practitioners in 'participatory design' (or 'co-design') have put forward a range of tested methodologies to include the views, needs, and interests of others (See, for example (Muller, 2007)). These methodologies have been adopted in computer science in sub-disciplines called 'human-centered computing' (see, for example (Brown et al., 1999; Jaimes, Sebe, & Gatica-Perez, 2006)) and 'social computing' (See, for example (Parameswaran & Whinston, 2007; Wang, Carley, Zeng, & Mao, 2007)). By actively and successfully involving potential research participants in the design of an experiment, researchers may be considered to be demonstrating a heightened sense of responsiveness.

Articulation of a new provision

The two challenges to the centrality and operationalisation of *autonomy* in the Common Rule, allow for the articulation of a new provision. Following from the previous section on human subjects, this subsection provides an answer to research question 3. The new provision introduces relational autonomy and the exploration of communication with distant participants.

§B Operationalising the central moral principle

(1) The REC may require the researcher to *characterise* their research relationships with human subjects, as identified in §A.1 and §A.2, in terms of their *vulnerability* and *dependency* on the actions and attitudes of the researcher. The REC may ask the researcher to justify whether particular characterisations of the research relationship do or do not apply (for example characterisations such as domination, oppression, exploitation).

(2) The REC may determine how the researcher informs human subjects about establishing a new research relationship.

- a) An informed consent procedure establishes an explicit relationship. The REC may determine the information provided to human subjects by the researcher and the channel for communication.
- b) A provision of information without the requirement to gain consent may be accompanied with further *safeguards*, such as an opt-out procedure.
- c) Human subjects are considered inherently vulnerable with regard to the researcher if the information requirement is waived by the REC, unless the REC determines that human subjects are not vulnerable.

(3) The REC may require the researcher to demonstrate whether the researcher identified the needs and interests of human subjects, how the needs and interests are understood, and to what extent the needs and interests of human subjects are interpreted and safeguarded in the research design. The REC may impose obligations to obtain additional knowledge about the needs and interests by collaborating with potential human subjects or representative organisations in the research design process.

Characterisation of a research relationship: A description of the outcome of a research relationship for participants, most notably human subjects.

Vulnerability: The extent to which human subjects are unable to withstand harm.

Dependency: The extent to which human subjects' situation is influenced by an others' attitude and action.

Safeguards: A measure taken to protect a human subject to prevent or reduce risks of harm.

Harm and privacy

This section analyses challenges from the conceptual framework to the Common Rule concepts of minimal risk of harm and privacy. These two sections are combined, because minimising the risk of harm to research participants and respecting their privacy are two closely related principles that are pivotal in the research ethics review procedures. One of the most prominently occurring harms in research is, in fact, a violation of human subjects' privacy. Similar to the previous sections, this section states the existing provisions for harm-avoidance and privacy in the Common Rule, analyses the challenges that have arisen throughout this thesis, and articulates a revised provision for the Common Rule.

Common Rule

A pivotal concept in the Common Rule is the risk of minimal harm that is inflicted on human subjects as a direct result of research studies. The law sets standards for the interpretation of the vague concept of harm and prescribes limitations to a potential causal chain of harms. The Common Rule defines the concept of minimal risk of harm in §46.102(i) as follows:

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

This formulation sets two standards. First, it reduces the harm concept to a moral accounting exercise, balancing the probability and magnitude of a harm occurring. Second, it defines the harm variable (in the calculation) to be the risks faced by an

average person in their daily life, assuming that all persons experience the same standard of risk. The standard “during the performance of routine physical or psychological examinations or tests” is not considered further in this analysis, because the law is assessed for a new type of research.

A similar type of moral accounting, where benefit and harms are put on a balance sheet, is prescribed again in §46.111(a)(2):

Risks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result. In evaluating risks and benefits, the IRB should consider only those risks and benefits that may result from the research (as distinguished from risks and benefits of therapies subjects would receive even if not participating in the research). The IRB should not consider possible long-range effects of applying knowledge gained in the research (for example, the possible effects of the research on public policy) as among those research risks that fall within the purview of its responsibility.

This provision limits the moral accounting exercise to only include the harms and benefits that result directly from the research. It explicitly requires IRBs to disregard the possible long-range effects of applying knowledge gained in the research. This limitation does not differentiate between intrinsic and extrinsic harms, per se, though it may not exclude responsibility for creating a causal sequence.

The concept of ‘minimal risk’ therefore consists of the following elements:

- A probability calculation of a qualitative variable (ie. harm and discomfort),

- A forecast or anticipation, though limited only to the direct effects of the study,
- A comparison to a 'daily life' standard, or,
- A comparison to 'routine physical or psychological examinations or tests',
- A balancing of risks of harm to the anticipated benefits to subjects and society.

The Common Rule is less clear about protections for human subjects' privacy. As a moral harm, privacy is therefore included in this section. There are two notable provisions where privacy is mentioned in the Common Rule.

First, to approve a research study, §46.111(7) requires an IRB to ascertain that "[w]hen appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of data". It is left unclear, however, when it would be appropriate for researchers to protect the privacy of research participants, and how to assess which provisions are adequate.

Second, the Common Rule offers several exemptions to research ethics review in §46.101(b), of which the most relevant concern the reuse of existing and publicly available information in sub-part (b), which is related to privacy:

Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

This provision exempts research that is specifically based on the reuse of existing data or that uses information from publicly available sources. At the time of drafting the Common Rule, lawmakers could not have foreseen the extensive capabilities of data analytics and resulting re-identification of individuals in seemingly anonymised datasets. Before the advent of digital technologies and networks, it was relatively obvious which information would be publicly accessible; the average person could either access it or not.

Challenges from conceptual framework

The first set of challenges question the appropriateness of the standards and delineations of the Common Rule provisions in the new sociotechnical environment of technical Internet research. The development of the conceptual framework in this thesis demonstrates the concept of harm, and the aspiration to minimise harm for all data subjects--including privacy violations--is elusive. Instead of a set of standards for compliance, this section proposes a duty of care for technical Internet researchers.

The second set of challenges concerns the responsibility for extrinsic and downstream harms that result for technical Internet research studies. The Common Rule excludes consideration of the longer-term effects of research. This section argues that a special duty of care for technical Internet researchers is appropriate for technical Internet research.

(1) From standards to duty of care

The standards set in the Common Rule, the prescribed moral accounting equation, and the limitations of reasoning about consequences indicate a strong compliance framework. The first challenge shifts the compliance framework to an interpretative

model. First, compliance framework is criticised from the perspective of the conceptual framework and technical Internet research generally. Then, following on from the previous section on autonomy, the source of morality is shifted from harm to vulnerability as part of the focus on relationships instead of individuals. Finally, a duty of care for technical Internet researchers is introduced.

Critique of the Common Rule equation for avoiding harm

The harm calculation prescribed in the Common Rule is not appropriate for technical Internet research. It requires a calculation of the probability and magnitude of harm or discomfort for human subjects anticipated as a result of participating in a research study. This formulation signals a universalist and abstract ethical (specifically consequentialist) reasoning, based on the assumption of proximate and particular human subjects. The shift in nature of human subjects has changed from the assumed proximate and particular other in the Common Rule, to a distant and generalised other in technical Internet research, complicates the prescribed harm equation. The distant human subjects may consist of a complex and dispersed web of others, with undefined or divergent sociopolitical parameters. This section describes why an accurate assessment of harm for data subjects is difficult (or likely impossible) to achieve with any meaningful accuracy, when using this equation.

The inputs for the calculation are a definition of harms, and a standard for a threshold. The conceptual framework raises the question whether harm can be meaningfully defined at all, which challenges the inclusion of harm as a standard for ethical evaluation. In lieu of an existing definition, the second chapter of the thesis found the following types of harm to be relevant in various research disciplines, among others: Physical harm or injury, discomfort, exposure to threats, moral harm,

emotional harm, distress, embarrassment, anxiety, violations of self-esteem, self-incrimination, reputational damage, inconvenience, stigmatisation, indignity, breaches of confidentiality, insurability. Several of these types of harm are found throughout the case studies of the thesis. However, most of these harms are not quantifiable, and therefore do not lend themselves to be included in a simple calculation. It will be hard, if not impossible, for a technical Internet researcher to assess the (quantitative) magnitude of these harms for their potentially diverse research population. And even if they were quantifiable, their selection, identification, or assessment would not be value-free, but be informed by the inherently limited contextual understanding of the individual researchers. The impact of various harms is likely to differ per individual human subject, especially when they are dispersed across a variety of sociopolitical contexts.

Further, the standards provided in the Common Rule to set a threshold for harm are inappropriate and irrelevant. The first standard requires a comparison to the risks of harms people face in their 'daily life', which can only be meaningfully operationalised with a homogeneous set of data subjects. The likelihood of a harm materialising depends on large part on the context within which people reside, and it is unlikely to be possible to quantify their magnitude. The materialisation and magnitude of harm to human subjects are therefore likely to be highly contextual. Imposing this standard would result in imposing a single standard of risk for a potentially hugely diverse research population. An accurate assessment of harm for data subjects is therefore difficult—if not impossible—to achieve with any meaningful accuracy. The second standard is only relevant for medical sciences. The prescribed comparison to a 'routine physical or psychological examinations or tests' is irrelevant for technical Internet research. This standard is not analysed further.

Finally, even if the calculation were a feasible and appropriate method to determine whether a standardised threshold of harm was overstepped, technical Internet researchers are unlikely to be equipped with the necessary skills to assess when a risk of harm is or is not 'minimal'. These largely depend on the identification and interpretation of social and political conditions that may differ from the researcher's lived experience. Further, to assess harms resulting from legal violations, researchers would (in effect) be required to conduct an international, possibly global legal diligence, to ensure they are not violating human subjects' rights.

Shift source of morality to vulnerability

The concept of 'harm', and therefore the Common Rule standard for 'no more than minimal risk of harm' do not appear to be appropriate sources of moral guidance for technical Internet researchers. Minimising potential risk of harm to human subjects equates to maximising harm-avoidance for a potentially undefinable, distant group of potentially implicated persons. This subsection analyses how the source of morality may be based on an ethics of care approach, whereby the characterisation of relationships between the researcher and their human subjects guide the imposition of responsibilities.

The shift from proximate and particular others to distant and generalised populations of data subjects requires a rethinking of the incentives researchers have to ensure a 'minimal risk of harm'. The original consequentialist moral accounting requires a hypothetical calculation of many non-quantitative variables, which become too wide-ranging and complex when the population of data subjects is potentially scattered throughout a large variety of sociopolitical contexts. The shift

towards an ethics of care approach underlying a legal research ethics framework invites a consideration of more relational approaches to the incentives of more powerful actors within contexts. Here, moral responsibility is derived from the accidental or purposeful embeddedness in particular social contexts, the power dynamics within these situations, and the vulnerability of those for whom the moral agent has a responsibility to care (Held, 2005).

In the conceptual framework and throughout the case studies of this thesis, human subjects are characterised as being vulnerable and dependent on the actions and decisions made by researchers. The previous section described the relationship between technical Internet researchers and their human subjects to be (potentially) hierarchical, dominating, exploitative, and oppressive. Human subjects fit the typical description of vulnerable populations, lacking a decisional capacity and to be likely to suffer greater harm. Here, harm is an outcome of the characterisation of the relationship. Approached in this way, it is possible to identify the types of harm that result from a particular attitude in a relationship

In the ethics of care literature, the required level of responsiveness is determined by the (negative) outcomes of the relationship. The previous section identified that responsiveness may result in a positive duty towards human subjects. This positive duty may be translated to a duty of care, which is discussed in the next section.

Shift from compliance towards duty of care

Duty of care is a legal concept with a specific meaning. It is the requirement and responsibility for people or organisations to avoid behaviours and to take reasonable measures to prevent harm caused to others. In general, jurisdictions apply a duty of

care to professionals or organisations is particular sectors. In case law, the duty of care is typically applied when a professional presents themselves as having the skills and knowledge needed to help another person with their need.

The duty of care is historically not derived from the ethics of care, though parallels and overlaps are obvious. For example, Bubeck argues that "[...] the best a carer can do is to minimise the harm she cannot prevent from occurring" (Bubeck, 1995, p. 202). Similarly, Kittay argues that moral harm is derived from failures to respond to needs and taking appropriate responsibility, instead of merely violating rights or causing generally harmful consequences (Kittay, 2011). Collins uses the same words that are used to describe a duty of care when she states that the ethics of care "instructs each agent to take both actions and omissions in order to ensure that they themselves do not harm others". She differentiates this duty from harm minimisation in moral accounting approaches, which "might instruct agents to cause harm to one person in order to minimise harm overall (Collins, 2015, p. 32). Finally, Nortvedt et. al connect the ethics of care (in the domain of health care) to duties of care, arguing that a relational view of the professionals' role obligations leads to particular duties of care towards their patients (Nortvedt et al., 2011).

The link between the ethics of care and duties of care become clearer when these professions are considered through the lens of the ethics of care. Typically, a duty of care is imposed on professionals such as doctors, lawyers, accountants, architects, to name only a few. Their commonality is that these professionals form relationships with people who are not (necessarily) knowledgeable of the specific tasks carried out by the professional on their behalf. The professionals thereby become powerful

actors within a hierarchical relationship to persons who are left relatively vulnerable to their decisions and attitudes.

The power imbalance between the professional and the people they serve is rebalanced by imposing a duty of care. By imposing a penalty on not acting in the best interest of clients, patients, or otherwise vulnerable persons, the professionals are disincentivised to prioritise their own interests. In medicine, for example, a doctor has a duty towards their patients, recommending treatments, informing them about alternatives, and to allow the patient some choice. Hereby, the doctor cannot be incentivised to, for example, prescribe medication from a pharmaceutical company that bribed them. The imposed duty of care ensure aligns the interests of the doctor with those of the patient (Spruit et al., 2016). Further, Nortvedt et. al note that, constructing the professionals' duties in this way, the doctor respects the uniqueness of each patient and takes their vulnerability into consideration (Nortvedt et al., 2011).

The conceptual framework reflects the description of professionals for the role of technical Internet researchers. They are professionals who have superior knowledge to their human subjects. Contrary to the professions described, the vulnerable persons in a technical Internet research relationship are typically unaware that they are, in fact, in a research relationship. However, even if the human subjects would be aware, they would typically lack the knowledge needed to understand and object meaningfully to research practices they deem to be in violation of research ethics or their own expectations. Imposing a duty of care on technical internet researchers disincentivises them from engaging in self-serving moral accounting, and to be actively responsive to the needs and interests of their implicated human subjects.

(2) Include extrinsic and downstream harm

The second challenge addresses the limitation of causal chains and consideration of future harms. The Common Rule limits an IRBs' ability to take into account the harms of a research beyond its direct impact in §46.111(a)(2):

In evaluating risks and benefits, the IRB should consider only those risks and benefits that may result from the research (as distinguished from risks and benefits of therapies subjects would receive even if not participating in the research). The IRB should not consider possible long-range effects of applying knowledge gained in the research (for example, the possible effects of the research on public policy) as among those research risks that fall within the purview of its responsibility.

This provision is somewhat unclear, because it excludes "possible long-range effects of applying knowledge gained in the research" in a research ethics assessment, though offers as an example only "the possible effects of the research on public policy". It may be argued that the provision also excludes considerations of precedent setting (e.g. enabling the reuse of password dumps) and downstream harms by third parties.

The impact of ignoring long(er)-range, indirect effects of research become abundantly clear in the second case study. Technical Internet users are able to make use of data subjects' devices and Internet connections and thereby make connections to devices owned and operated by third parties. These further connections create the perception by third parties that the data subjects themselves are requesting

information. The researchers thereby attribute actions to data subjects without their knowledge.

As a result, the researchers initiate a causal chain whereby their data subjects may experience harm, if the third parties decide to act on their perception. Resulting downstream harm can occur, depending on the nature of the attributed action (ie. whether it is unlawful or politically / socially sensitive in the data subject's context), the capacity of the third party to interpret the request, the political expediency, and the legal directive to investigate or enforce against the action. Therefore, the level of foreseeable downstream harms should also be taken into account in the assessment of the risk of harm, and the moral loss to the discipline as a whole.

In her assessment of approaches to global security issues, Robinson describes that "people experience their moral lives in the context of webs of relationships with individuals and groups of particular others and that one of the main tasks of moral inquiry is to think about how care and responsibilities for care are distributed both within and across societies" (Robinson, 2011, p. 31). If care is explicitly limited to only the direct consequences of research, as prescribed by the Common Rule, responsibility and care are not distributed in a way that reflects the moral concerns that result from technical Internet research. Robinson suggests further that an ethics of care approach "actually offers us resources for critical evaluation of the nature of the responsibilities of care leading to revisions in resolutions of important questions regarding distribution, resourcing, and recognition of our moral responsibilities to others" (p. 31).

Therefore, this challenge simply suggests that moral concern should not be limited for technical Internet researchers, but their actions and attitudes should take into account future downstream harms for human subjects, the reputation of research institutions, and their effect of society. Researchers and research ethics boards should work together to develop responsive attitudes that identify, understand, and interpret future harms for others, and to justify their experimental design with this knowledge.

Articulation of a new provision

The Common Rule provides a set of inappropriate standards and an over-simplified equation to determine whether a proposed research inflicts harm on human subjects. Research ethics review would be better served by a legal framework that supports researchers and reviewers in making decisions when the risks of harms and societal benefits conflict. Responsiveness becomes increasingly important when values or harms are not quantifiable, though a harm is foreseeable. The provision below incorporates the challenges of this section to communicate to researchers and research ethics reviewers the need for responsiveness and offers an interpretable framework for decision-making.

§C Responsiveness to harm

(1) The researcher must identify foreseeable and potential harms to their human subjects. These include, but are not limited to, direct (intrinsic) harms that result from the research (e.g. financial loss, violation of privacy) and indirect (extrinsic) harms (downstream harms caused by third parties as a result of research). The REC evaluates whether the identified harms are sufficiently comprehensive.

(2) The researcher must describe how the various harms are likely to impact human subjects, and how these harms are mitigated or reduced in the research design. The REC supports the researcher to identify which external groups may be consulted to verify whether the research design is responsive to the needs and interests of potential human subjects.

(3) The researcher and the REC collaborate to develop the study with the watchfulness, attention, caution, and prudence that a reasonable person in the circumstances would use to avoid unnecessary harms. If the researcher's actions do not meet this standard of care, then the research study is considered negligent and does not qualify for research ethics approval.

(7) Conclusion

This chapter marks the end of the thesis. It reflects briefly on the research process and reformulates the answers to research questions. Further, it suggests some practical contributions of the thesis, and new avenues for future research that have become apparent.

I decided to write a thesis challenging the research ethics framework for technical Internet research after seeing how clearly problematic research proposals were allowed to proceed after a research ethics review. I noticed a discrepancy about how computer scientists and technical Internet researchers posited self-serving ethical arguments for their research, and research ethics committees who missed the technical knowledge to challenge the researchers meaningfully. As the case study papers in this thesis demonstrate, ethics committees have allowed studies to commence that harm their implicated human subjects in a variety of ways, set undesirable precedents, set up human subjects for downstream harms, and can be potentially negligent.

However, the problem was not that technical Internet researchers were malevolent. Quite the opposite: from my personal experience I found them to be a largely benign community, though where an ‘end justify the means’ approach was dominant. This led to supposedly positive research outcomes using any tools and techniques available without much regard for the external and collateral effects of their studies. As long as a particular goal was achieved, it was considered “too bad” that people may be temporarily implicated in research studies.

Similarly, members of ethics boards are typically very thoughtful people. However, the tools they have been provided with to assess research proposals, stem from unrelated research disciplines and from another era. This thesis focused on the US Common Rule, though the same can be said about other research ethics policies in other jurisdictions. The biomedical origins of the Common Rule were obvious in the assessment of the case studies in this thesis. Not only does the law contain a range of biomedical-specific terms (e.g. “venepuncture” and “patients”), but the assumptions underlying the law also indicate that the legislator had a doctor – patient relationship in mind when drafting the law.

James Moor called this situation a policy vacuum, where technology has overtaken existing law and policy in a way that the rules can no longer be satisfactorily applied. For example, the law assumes a proximate and particular research relationship, while implicated human subjects in technical Internet research are distant and generalised. The researcher typically does not know who is providing the research data, and the human subjects are not aware they are being experimented on. This difference undercuts several provisions in the Common Rule, such as the primacy of informed consent, standards for harm, and the awareness that humans are, in fact, implicated in the study and not only some distant devices.

I chose to use the structured case methodology as initially described by Carroll and Swatman to build a constructive framework for analysing case studies in technical Internet research ethics. Their methodology stems from the neighbouring field of information systems research and it allows for a thorough, literature-based scrutiny of case studies. I decided to construct this thesis on a foundation of research ethics law and scholarship, before moving to the lessons learned in social science Internet

research. Combined, the knowledge gathered from these fields provided the first conceptual framework to assess case studies. The conceptual framework was enriched by each case study, adding new ideas and concepts to the tables at the end of each chapter. This culminated in a final analytical chapter, where the conceptual framework formed the basis for a more in-depth analysis of the concepts in the Common Rule.

It became apparent that the ethics of care would serve as a suitable ethical framework to underpin the thesis. I found computer scientists to have a consequentialist mindset (“the ends justify the means”) and the Common Rule to contain a mixture of deontological ideas (e.g. informed consent) and consequentialist standard setting (e.g. moral accounting to quantify harm and benefit). Since this system was clearly not working for the evaluation of technical Internet research, I decided to experiment with a new underlying ethical framework. Given the myriad of relations a researcher has—to their human subjects, colleagues, and society as a whole—the ethics of care became an obvious framework to apply.

I initially expected this research approach to yield some interesting insights into how the Common Rule could be strengthened as a legal framework that could also guide technical Internet research ethics. However, as the study progressed, I realized that the underlying assumptions of the Common Rule were not appropriate for the features of technical Internet research. This realisation led me to propose some new legal provisions for technical Internet research ethics review specifically. The next section describes them in a more detail.

Answers to research questions

This thesis was driven by three research questions. The first question is the main focus of the thesis, which was supported by two other questions. This section answers the questions briefly

(1) How ought technical Internet research with unsuspecting data subjects be assessed from the perspective of research ethics?

This thesis proposes three new legal provisions that aim to develop research ethics review in two new directions.

First, the provisions require researchers to adopt a responsive approach to the persons and groups who they are in a relationship with. For example, this includes human subjects who are implicated or involved in the research study in a variety of ways, the researcher's colleagues in their institutions and research disciplines, and wider society. The emphasis is on developing a positive relational attitude with the human subjects, who are most directly affected by the researchers' attitudes. Positive relationships with others will most likely follow from a positive research relationship to human subjects. To this end, the first provision allows researchers to identify whether people are implicated in a study when the assumption is that only technical devices are included. The second provision shifts the burden of ethical reflection more firmly on the researcher, instead of the potential participant (e.g. through an informed consent procedure). Finally, the third provision develops a duty of care for the researcher, which allows implicated human subjects to claim negligence when they have been harmed by a research study.

Second, the position of a research ethics committee (known as an institutional review board in the Common Rule) is changed throughout these new provisions. One of the effects of the new provisions is that a research ethics committee should play a more active role in holding researchers to account and working with them to unearth the effects of their research in a complex sociotechnical reality. This idea stems from Visse et. al, who argue that ethics committees should not merely execute a decision tree full of concepts and categories to mechanically assess whether a project is appropriate. Instead, a research ethics committee should be a *responsive evaluator* who works with the researchers to identify and understand the interests and needs of others (i.e. implicated human subjects), and to interpret their concerns into the research design (Visse et al., 2015). The research ethics committee should challenge the researchers constructively and help justify the design choices that were made. The provisions developed in this thesis provides topics and methods for the research ethics committee to discuss and develop with the researchers.

(2) Which concepts and principles of research ethics frameworks are challenged specifically by technical Internet research with unsuspecting data subjects?

The new provisions are the result of concepts and principles that were identified as inappropriate or insufficient throughout the case studies. The ‘human subject’ concept is too narrowly defined in the Common Rule, and rests on the assumption that researchers know who their research subjects are. Further, the guiding principle of autonomy is irrelevant when a large part of the research in a discipline is conducted without the knowledge or consent of the research participants. Similarly, standards of harm and moral accounting equations are not a useful guide when research subjects are not homogenous groups (as assumed by the Common Rule),

but potentially dispersed around the world in a variety of socio-political contexts. Finally, privacy is a poorly defined concept in the Common Rule, though privacy violations are likely the most common form of harm for research subjects.

(3) How might these concepts be reinterpreted for the purposes of research ethics evaluation of technical internet research when considered through the appropriate ethical framework?

The challenged concepts and principles of the Common Rule are based on consequentialist and deontological ethical frameworks. To meaningfully challenge the concepts, I chose to first find the appropriate ethical framework. The ethics of care served this purpose well. Its focus on positive relationships instead of abstract moral accounting or focus adherence to autonomy is a primary value, offered an interesting lens to assess the challenges to the concepts and principles identified in the second research question. The reinterpretations of the concepts are described at length in the sixth chapter of this thesis. For the purpose of this concluding chapter, it is sufficient to state the assumptions for the reinterpretation. First, human subjects are typically unaware and at a distance from researchers. This distance principle typically results in researchers and ethics boards failing to adequately account for the interests and needs of implicated human subjects. Therefore, the new provisions which are based on an ethics of care framework, incentivise researchers and research ethics committees to actively identify and understand the needs to implicated human subjects and to interpret and justify them in the final research design. Therefore, simple standards and equations are not sufficient. Researchers must understand the impact of their experimental design in the context of their implicated human subjects.

Practical contributions

In the introductory chapter, the aim of this thesis is described as assessing “the extent to which the research ethics frameworks imposed on technical Internet researchers in academic institutions—in disciplines such as computer science and network engineering—are fit-for-purpose, whether they yield morally satisfactory outcomes, and, if not, how they should be reinterpreted or reviewed.” The case studies, the literature-based scrutiny, and the newly proposed provisions in the sixth chapter contribute to this aim. This section describes the practical outcomes for the four audiences for whom these findings are relevant.

Research Ethics Committees (and Institutional Review Boards)

Research ethics committees are tied to the rigid mechanisms and standards set in the Common Rule. They cannot deviate from the concepts, principles, and structure of the law in their assessment of research. For a new sociotechnical field such as technical Internet research, this means that members of a research ethics review committee must apply somewhat inappropriate or unsuitable concepts to new research proposals.

This thesis offers a new approach, where the research ethics committee acts as a responsive evaluator. The provisions developed in this thesis are purposefully less rigid and a little vaguer than the Common Rule, which allows the research ethics committee discretion to help researchers identify, understand and interpret the needs of their research subjects. In line with a shift in legal thinking towards more communicative approaches, the new provisions communicate which values are important, and suggest some mechanisms to operationalise them.

Technical Internet researchers

Technical Internet researchers rarely have a thorough understanding of social dynamics of the contexts within which their research takes place. The rigid structure of the Common Rule provides researchers with a system to assess their own ethical decision making. Sadly, the Common Rule provides a false impression that researchers are engaged in ethics and it offers a series of out-dated and inappropriate standards and concepts.

The new provisions in this thesis provide more appropriate guidance for scholars designing and conducting technical Internet research when assessing the research ethics requirements of their studies. Instead of compliance with standards and equations, the provisions incentivise researchers to gather social indicators about the needs and interests of their human subjects, which in turn should influence the research design. Further, it provides researchers with a research ethics committee that is tasked with thinking along constructively with the researcher, and challenging their assumptions.

Policymakers

The practical contribution for policymakers is two-fold. First, this thesis has shown the current approach to research ethics law for technical Internet research to yield inappropriate outcomes and in need of review. The outcome of the thesis is a law that communicates important social values, mechanisms to operationalise these values, with the intent to align the values of researchers and human subjects. Second, the thesis demonstrated an approach or a methodology to conduct such a legislative review. When it is apparent that the sociotechnical constituency of the regulated environment has shifted dramatically (i.e. from medical contexts to global information networks), a legislative review requires a deeper evaluation that

addresses the underlying assumptions and ethical frameworks. This thesis is an attempt to demonstrate that this is possible and necessary for information laws that govern society's evolved interaction with data.

Ethics of care scholarship

Finally, this thesis contributes to the ethics of care literature. It was not the aim of this thesis to extend the theories and concepts of the ethics of care, but merely to apply them. However, by applying considerations about research relationships, responsiveness in the realm of technical Internet research, and a review of specific concepts of the Common Rule, I hope that this thesis will have contributed in the applied sense. At the time of writing, I could only find a handful of papers that used the ethics of care framework to address emerging issues as a result of the Internet and digitisation of social interactions. This thesis is another contribution in that space of the ethics of care.

Future research

This thesis is interdisciplinary in nature, which in practice means that it doesn't fit neatly into one specific discipline. It uses concepts from philosophy and law and applies them computer science topics. Therefore, no discipline was fully leading throughout the thesis, and decision had to be made about the depth of analysis of the various disciplines. Given the space and time constraints on a DPhil thesis, some aspects could have been developed further. Fortunately, this thesis builds a fertile scholarly ground for further investigations. I propose three distinct areas that are ripe for further research.

New areas of computer science

Computer science is an ever-evolving discipline that contributes to, but also follows ongoing technical innovation. This thesis is limited to two cases that are exemplary of technical Internet research, though the domain of computer science is broader than that. For example, machine learning has developed impressively over the last few years. The field is confronted by myriad ethical issues and questions. As a primarily research-driven field, machine learning could benefit greatly from a new research ethics framework that takes into account the various stages where researchers make decisions that may have a direct or indirect social consequences, for their human subjects or others who may later be susceptible to the complex algorithms.

Negligence and a duty of care

The legal concepts of negligence and duties of care can be usefully applied to researchers to align their interests with research subjects. However, these legal constructs are much more complex and intellectually deep than presented in this thesis. It would warrant a new DPhil project to interpret the elements of case law from various jurisdictions to the professional duties of a technical Internet researcher, or to the broader group of professionals in computer science.

Review other information laws using this methodology

Finally, I would like to recommend that all law relating to digital information are reviewed thoroughly using the methodology of this thesis. Most information laws that are in the books today, were developed in an era when information was predominantly printed on physical carriers (e.g. newspapers, books, vinyl, magnetic discs), broadcasted, or stored on individual terminal computers that may or may not have been connected to a local network. The formulation of these laws made sense

for the technologies they were regulating. However, the Internet has dramatically changed how information is created, disseminated, collected, stored and accessed. Further, digital information and globally interconnected information networks have significantly changed how people interact with information. New ethical issues emerge frequently, which could not or were not foreseen by legislators. Therefore, it is prudent to review all laws that govern information and data in whether it achieves its aims, whether the underlying philosophical frameworks are appropriate, and whether the law should be amended, rewritten, or discarded. In the case of research ethics law for technical Internet research, this thesis recommends developing a new research ethics law or policy, which replaces or supplements the Common Rule for this specific discipline.

References

- A29WP. (2014). ARTICLE 29 DATA PROTECTION WORKING PARTY Opinion 05 / 2014 on Anonymisation Techniques Adopted on 10 April 2014.
- Abbate, J. (2000). *Inventing the Internet*. MIT Press.
- Aby, S. H., & Kuhn, J. C. (2000). *Academic freedom: A guide to the literature* ((20)). Greenwood Publishing Group.
- Aceto, G., & Pescapé, A. (2015). Internet censorship detection: A survey. *Computer Networks*, 83, 381-421.
- Acquisti, A., & Grossklags, J. (2004). Privacy attitudes and privacy behavior. In *Economics of information security* (pp. 165-178). Springer.
- Agre, P. E. (2003). P2P and the promise of internet equality. 46(2), 39-42.
- Allen, C. (1996). What's wrong with the 'golden rule'? Conundrums of conducting ethical research in cyberspace. *The Information Society*, 12(2), 175-188.
- Altman, I. (1975). *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding*.
- Anderson, C. (2008). The end of theory: the data deluge makes the scientific method obsolete. *Wired magazine*, 16(7).
- Anderson, G., & Arsenault, N. (2005). *Fundamentals of Educational Research* (Falmer Press Teachers' Library).
- Anonymous. (2012). Internet Census 2012.
- Anscombe, G. E. M. (1958). Modern moral philosophy. *Philosophy*, 33(124), 1-19.
- APA. (2002). APA Ethical Guidelines. *The American psychologist*, 57(12), 1060-1073.
- Arnold, T. W. (1937). The symbols of government.
- Ashford, N., Barrett, K., Bernstein, A., Costanza, R., Costner, P., Cranor, C. et al. (1998). *Wingspread statement on the precautionary principle*. Available at: <https://kundaliniandcelltowers.com/precautionary-principle-wingspread-statement.pdf>.
- Austin, J. (1863). *The Province of Jurisprudence Determined: Being the First Part of a Series of Lectures on Jurisprudence, Or, the Philosophy of Positive Law* (2). J. Murray.
- Aycock, J., Buchanan, E., Dexter, S., & Dittrich, D. (2011). *Human subjects, agents, or bots: Current issues in ethics and computer security research*. Springer, Berlin, Heidelberg.
- Azua, M. (2009). *The social factor: Innovate, ignite, and win through mass collaboration and social networking*. Pearson Education.
- Baggini, J., & Fosl, P. S. (2007). *The ethics toolkit: A compendium of ethical concepts and methods*.
- Bailey, M., Kenneally, E., & Dittrich, D. (2012). A refined ethical impact assessment tool and a case study of its application., 12.
- Ballot, E., Montreuil, B., & Meller, R. (2014). *The Physical Internet: The Network of Logistics Networks*. La Documentation Française.
- Barnes, M., & Brannelly, T. (2015). *Ethics of care: Critical advances in international perspective*. Policy Press.
- Bassett, E. H., & O'Riordan, K. (2002). Ethics of Internet research: Contesting the human subjects research model. *Ethics and Information Technology*, 4(3), 233-247.
- Baumrind, D. (1985). Research using intentional deception: Ethical issues revisited. *American Psychologist*, 40(2), 165.
- Baumrind, D. (2015). When subjects become objects: The lies behind the Milgram legend. *Theory & Psychology*, 25(5), 690-696.
- Beck, U. (1992). *Risk Society: Towards a New Modernity*. (73(1)). SAGE Publications.
- Beecher, H. K. (1966). Ethics and clinical research. In *Biomedical ethics and the law* (pp. 215-227). Springer.

- Benatar, S. R., & Singer, P. A. (2000). A new look at international research ethics. (*BMJ*) *British Medical Journal*, 321(7264), 824.
- Benčin, R., Šumič-Riha, J., & Riha, R. (2015). Ethics Assessment in Different Fields: Humanities. *Stakeholders Acting Together on the Ethical Impact Assessment of Research and Innovation - SATORI, Deliverable 1.1, Annex 2.e*.
- Bennett, C. J. (1991). Computers, personal data, and theories of technology: Comparative approaches to privacy protection in the 1990s. *Science, Technology, & Human Values*, 16(1), 51-69.
- Berman, J., & Weitzner, D. J. (1997). Technology and democracy. *Social Research*, 1313-1319.
- Berners-Lee, T., & Fischetti, M. (2001). *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*.
- Berry, D. M. (2004). Internet research: privacy, ethics and alienation: an open source approach. *Internet research*, 14(4), 323-332.
- Boddy, J., Neumann, T., Jennings, S., Morrow, V., Alderson, P., Rees, R. et al. (2010). The research ethics guidebook: a resource for social scientists. *The research ethics guidebook: a resource for social scientists*.
- Boehlefeld, S. P. (1996). Doing the right thing: Ethical cyberspace research. *The Information Society*, 12(2), 141-152.
- Boele-Woelki, K., Fisher, C., Pape, H., & Francisco, J. S. (2019). Securing Academic Freedom through Networks. *Angewandte Chemie International Edition*, 58(25), 8246-8248.
- Boellstorff, T. (2012). Rethinking Digital Anthropology. In H. A. Horst & D. Miller (Eds.), *Digital Anthropology*. London: Berg.
- Bondolfi, A. (2000). Ethics, Law and Legislation: The Institutionalisation of Moral Reflection. *Ethical Theory and Moral Practice*, 3(1), 27-37.
- Bonneau, J. (2012). *The science of guessing: analyzing an anonymized corpus of 70 million passwords*.
- Bonneau, J., Herley, C., & Stajano, F. M. (2014). Passwords and the evolution of imperfect authentication.
- Bourdieu, P. (1988). *Homo academicus*. Palo Alto, CA: Stanford University Press.
- boyd, D. (2007). Why youth (heart) social network sites: The role of networked publics in teenage social life., 119-142.
- Bracken-Roche, D., Bell, E., Macdonald, M. E., & Racine, E. (2017). The concept of 'vulnerability' in research ethics: an in-depth analysis of policies and guidelines. *Health Research Policy and Systems*, 15(1), 8.
- Bradley, B. (2012). Doing Away with Harm. *Philosophy and Phenomenological Research*, 85(2), 390-412.
- Brannelly, T. (2018). An ethics of care research manifesto. *International Journal of Care and Caring*, 2(3), 367-378.
- Society, B. P. (2017). Ethics Guidelines for Internet-mediated Research., 23.
- Brown, I. (2013). Future Identities: Changing identities in the UK ,À the next 10 years.
- Brown, J. R., van Dam, A., Earnshaw, R., Encarnação, J., Guedj, R., Preece, J. et al. (1999). Human-centered computing, online communities, and virtual environments. *IEEE Computer Graphics and Applications*, 19(6), 70-74.
- Bruckman, A. (2002). Studying the amateur artist: A perspective on disguising data collected in human subjects research on the Internet. *Ethics and Information Technology*, 4(3), 217-231.
- Bu, T., & Towsley, D. (2002). *On distinguishing between Internet power law topology generators*.
- Bubeck, D. E. (1995). Care, gender, and justice.

- Buchanan, E. A. (2011). *Internet research ethics: Past, present, and future* (11).
- Buchanan, E. A., & Ess, C. M. (2009). Internet research ethics and the institutional review board: Current practices and issues. *ACM SIGCAS Computers and Society*, 39(3), 43-49.
- Buchanan, E. A., & Ess, C. (2008). *Internet research ethics: The field and its critical issues*.
- Buchanan, E. A., & Ess, C. M. (2012). *Internet Research Ethics*. Metaphysics Research Lab, Stanford University.
- Buchanan, E., Aycock, J., Dexter, S., Dittrich, D., & Hvizdak, E. (2011). Computer science security research and human subjects: Emerging considerations for research ethics boards. *Journal of Empirical Research on Human Research Ethics*, 6(2), 71-83.
- Burnett, S., & Feamster, N. (2015). *Encore: Lightweight measurement of web censorship with cross-origin requests*.
- Byers, J. W. (2015). *Encore: Lightweight Measurement of Web Censorship with Cross-Origin Requests—Public Review*.
- Capron, A. M. (2014). Subjects, Participants, and Partners: What Are the Implications for Research as the Role of Informed Consent Evolves. *Human Subjects Research Regulation: Perspectives on the Future*, 154-168.
- Capurro, R., & Pingel, C. (2002). Ethical issues of online communication research. *Ethics and Information Technology*, 4(3), 189-194.
- Carpenter, K., & Dittrich, D. (2011). Bridging the distance: removing the technology buffer and seeking consistent ethical analysis in computer security research., 29.
- Carroll, J. M., & Swatman, P. A. (2000). Structured-case: a methodological framework for building theory in information systems research. *European journal of information systems*, 9(4), 235-242.
- Celsus, C. (47 AD). *De Medicina*.
<http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Celsus/home.html>.
- Prevention, C. F. D. C. A. (2003). HIPAA privacy rule and public health. Guidance from CDC and the U.S. Department of Health and Human Services. *MMWR: Morbidity and mortality weekly report*, 52(1-17), 19.
- Chadwick, R., Ten Have, H., & Meslin, E. M. (2011). *The SAGE Handbook of Health Care Ethics*. Sage Publications.
- Chaudet, C., Fleury, E., Lassous, I. G., Rivano, H., & Voge, M.-E. (2005). Optimal positioning of active and passive monitoring devices. Proceedings from Proceedings of the 2005 ACM conference on Emerging network experiment and technology.
- Christman, J. (2008). Autonomy in moral and political philosophy. *Stanford encyclopedia of philosophy*.
- Claffy, K. C. (2000). Measuring the internet. *IEEE Internet Computing*, 4(1), 73-75.
- Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, 58(1), 7-19.
- Cobb, W. M. (1973). The Tuskegee syphilis study. *Journal of the National Medical Association*, 65(4), 345.
- Cohen, I. G., & Lynch, H. F. (2014). *Human subjects research regulation: Perspectives on the future*. MIT Press.
- Cohen, J. E. (2000). Privacy, ideology, and technology: A response to Jeffrey Rosen. *Geo. LJ*, 89, 2029.
- Cohen, J. E. (2012). What Privacy Is For. *Harv. L. Rev.*, 126, 1904.
- Colledge, F., Elger, B., & Howard, H. C. (2013). A review of the barriers to sharing in biobanking. *Biopreservation and biobanking*, 11(6), 339-346.
- Collingridge, D. (1980). *The Social Control of Technology*.
- Collins, S. (2015). *The core of care ethics*. Springer.

- Cotterrell, R. (1997). *Law's community: legal theory in sociological perspective*. Oxford University Press.
- Europe, C. O. (2012). Guide for Research Ethics Committee Members Steering Committee on Bioethics.
- Council of the European Union, E. P. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L119, 1-17.
- Crandall, J. R., Crete-Nishihata, M., & Knockel, J. (2015). Forgive Us our SYNs: Technical and Ethical Considerations for Measuring Internet Filtering., 7.
- Crawford, K., & Schultz, J. (2013). Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms.
- Crovella, M., & Krishnamurthy, B. (2006). *Internet measurement: infrastructure, traffic and applications*. John Wiley & Sons, Inc.
- Cutcliffe, J. R., & Ramcharan, P. (2002). Leveling the playing field? Exploring the merits of the ethics-as-process approach for judging qualitative research proposals. *Qualitative Health Research*, 12(7), 1000-1010.
- Das, A., Bonneau, J., Caesar, M., Borisov, N., & Wang, X. (2014). *The Tangled Web of Password Reuse*.
- Davis, M. (2002). Do the professional ethics of chemists and engineers differ.
- De Carnavalet, X. D. C., & Mannan, M. (2014). *From Very Weak to Very Strong: Analyzing Password-Strength Meters*. Proceedings from NDSS.
- Deacon, A., & Williams, F. (2004). Introduction: Themed section on care, values and the future of welfare. *Social Policy and Society*, 3(4), 385-390.
- Deakin, C. T., Alexander, I. E., & Kerridge, I. (2009). Accepting Risk in Clinical Research: Is the Gene Therapy Field Becoming Too Risk-averse. *Molecular Therapy*, 17(11), 1842-1848.
- Denning, P. J., Comer, D. E., Gries, D., Mulder, M. C., Tucker, A., Turner, A. J. et al. (1989). Computing as a discipline. *Computer*, 22(2), 63-70.
- Devlin, P. B. (1959). *The Enforcement of Morals*.
- Dingwall, R. (2006). Volume 1 Issue 1 November 2006 Liverpool University Editorial Team : Helen Bromley, Sara Medical Sociology online University of Liverpool A free access publication of the British Sociological Association. 1, 51-58.
- Dittrich, D., Kenneally, E., & Bailey, M. (2012). Applying ethical principles to information and communication technology research: A companion to the Department of Homeland Security Menlo Report.
- Dittrich, D., & Kenneally, E. (2012). Applying Ethical Principles to Information and Communication Technology Research : A Companion to the Department of Homeland Security Menlo Report.
- Doty, N. (2016). Mitigating browser fingerprinting in web specifications.
- Dove, E. S., Kelly, S. E., Lucivero, F., Machirori, M., Dheensa, S., & Prainsack, B. (2017). Beyond individualism: Is there a place for relational autonomy in clinical practice and research. *Clinical ethics*, 12(3), 150-165.
- Dutton, W. H., & Piper, T. (2010). *The Politics of Privacy, Confidentiality, and Ethics: Opening Research Methods*.
- Dutton, W. (2013). Internet studies: The foundations of a transformative field.
- Edelson, P. J. (2004). Henry K. Beecher and Maurice Pappworth: honor in the development of the ethics of human experimentation. 2, 219-233.
- Edwards, S. J. L., Ashcroft, R., & Kirchin, S. (2004). Research ethics committees: differences and moral judgement. *Bioethics*, 18(5), 408-427.

- Egelman, S., Bonneau, J., Chiasson, S., Dittrich, D., & Schechter, S. (2012). *It's not stealing if you need it: A panel on the ethics of performing research using public data of illicit origin.*
- Egelman, S., Sotirakopoulos, A., Muslukhov, I., Beznosov, K., & Herley, C. (2013). *Does my password go up to eleven?: the impact of password meters on password selection.*
- Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of management journal*, 50(1), 25-32.
- El Emam, K. (2010). Risk-based de-identification of health data. *IEEE Security & Privacy*, 8(3), 64-67.
- Elahi, T., & Goldberg, I. (2012). CORDON—a taxonomy of internet censorship resistance strategies. *University of Waterloo CACR*, 33.
- Elger, B. S., & Caplan, A. L. (2006). Consent and anonymization in research involving biobanks: differing terms and norms present serious barriers to an international framework. *EMBO reports*, 7, 661-666.
- Elgesem, D. (2002). What is special about the ethical issues in online research. *Ethics and Information Technology*, 4(3), 195-203.
- Elms, A., Erwin, E., Gendin, S., & Kleiman, L. (1994). Keeping deception honest: Justifying conditions for social scientific research stratagems. In (pp. 121-140). New York, NY: Garland.
- Engster, D. (2015). *Justice, care, and the welfare state*. Oxford University Press, USA.
- Engster, D., & Hamington, M. (2015). *Care ethics and political theory*. Oxford: Oxford University Press.
- Ensafi, R., Knockel, J., & Alexander..., G. (2014). Detecting intentional packet drops on the Internet via TCP/IP side channels. ... *Conference on Passive ...*
- EPIC. (2007). *Privacy and Human Rights Report 2006: An International Survey of Privacy Laws and Developments*. Electronic Privacy Information Center; 2006 edition (September 3, 2007).
- Eriksson, B., Dasarathy, G., Barford, P., & Nowak, R. (2012). Efficient network tomography for internet topology discovery. *IEEE/ACM Transactions on Networking (TON)*, 20(3), 931-943.
- Eriksson, S., & Helgesson, G. (2005). Potential harms, anonymization, and the right to withdraw consent to biobank research. *European Journal of Human Genetics*, 13(9), 1071.
- Ess, C. (2002). Ethical decision-making and Internet research: Recommendations from the AoIR ethics working committee. *Issues and controversies*, 27-44.
- Ess, C. (2006). Ethical pluralism and global information ethics. 8(4), 215-226.
- Commission, E. U. (2012). *Responsible Research and Innovation: Europe's Ability to Respond to Societal Challenges*. Brussels: Publications Office of the European Union.
- Evans, J. H. (2000). A Sociological Account of the Growth of Principlism. *Hastings Center Report*, 30(5), 31-39.
- Eynon, R., Schroeder, R., & Fry, J. (2009). New techniques in online research: Challenges for research ethics. *Twenty-First Century Society*, 4(2), 187-199.
- Eynon, R., Fry, J., & Schroeder, R. (2008). The ethics of internet research. *Sage internet research methods*, 23-41.
- Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. *BMJ*, 323(7321), 1103-1105.
- Faden, R. R., & Beauchamp, T. L. (1986). A history and theory of informed consent., 261-262.
- Fairchild, A. L., & Bayer, R. (1999). Uses and abuses of Tuskegee. *Science*, 284(5416), 919-921.

- Feldman, H. L. (2012). What's Right about the "Medical Model" in Human Subjects Research Regulation.
- Fieser, J. (2018). Ethics. <http://www.iep.utm.edu/ethics/>.
- Fiesler, C., & Proferes, N. (2018). "Participant" Perceptions of Twitter Research Ethics. *Social Media+ Society*, 4(1).
- Fineman, M. A. (2010). The Vulnerable Subject : Anchoring Equality in the Human Condition. 20(1), 171-191.
- Flanagin, A. J., Flanagin, C., & Flanagin, J. (2010). Technical code and the social construction of the internet. *New Media & Society*, 12(2), 179-196.
- Florencio, D., & Herley, C. (2007). *A large-scale study of web password habits*.
- Forell, C., & Sortun, A. (2008). The tort of betrayal of trust. *U. Mich. JL Reform*, 42, 557.
- Frankel, M. S., & Siang, S. (1999). Ethical and legal aspects of human subjects research on the Internet. *AAAS online*.
- Friedman, M. (2015). Privacy, Surveillance, and Care Ethics. *Care Ethics and Political Theory*, 108-126.
- Galligan, D. (2006). *Law in modern society*. OUP Oxford.
- Garfinkel, S. (2000). *Database nation: The death of privacy in the 21st century*. " O'Reilly Media, Inc."
- Gelinas, L., Wertheimer, A., & Miller, F. G. (2016). When and why is research without consent permissible. *Hastings Center Report*, 46(2), 35-43.
- Gilligan, C. (1977). In a Different Voice: Women's conceptions of self and of morality. *Harvard educational review*, 47(4), 481-517.
- Gómez-Vírseda, C., de Maeseneer, Y., & Gastmans, C. (2019). Relational autonomy: what does it mean and how is it used in end-of-life care? A systematic review of argument-based ethics literature. *BMC medical ethics*, 20(1), 76.
- Goodin, R. E. (1986). *Protecting the Vulnerable: A Re-Analysis of Our Social Responsibilities*. University of Chicago Press, 1986.
- Gotterbarn, D. W., Bruckman, A., Flick, C., Miller, K., & Wolf, M. J. (2018). ACM code of ethics: a guide for positive action.
- Green, L. (2013). The Morality in Law.
- Grimmelmann, J. (2015). The law and ethics of experiments on social media users. *Telecomm. & High Tech. L*, 13, 219.
- Grinter, R. E., Edwards, W. K., Newman, M. W., & Ducheneaut, N. (2005). The work to make a home network work.
- Grunwald, A. (2011). Responsible innovation: Bringing together technology assessment, applied ethics, and STS research. *Enterprise and Work Innovation Studies*, 31, 10-19.
- Guba, E., & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Thousand Oaks, CA: SAGE Publications, Inc.
- Gummadi, K. P., Saroiu, S., & Gribble, S. D. (2002). *King: Estimating latency between arbitrary internet end hosts*.
- Gurzawska, A., & Benčin, R. (2015). Ethics assessment in different fields: Social sciences. *Stakeholders Acting Together on the Ethical Impact Assessment of Research and Innovation - SATORI, Deliverable 1.1, Annex 2.d*.
- Haker, H. (2013). *Autonomy and care in medicine*.
- Hall, T., & Flynn, V. (2001). Ethical issues in software engineering research: a survey of current practice. *Empirical Software Engineering*, 6(4), 305-317.
- Hanamsagar, A., Woo, S. S., Kanich, C., & Mirkovic, J. (2018). *Leveraging semantic transformation to investigate password habits and their causes*.
- Hart, H. L. A., & Green, L. (2012). *The concept of law*. Oxford University Press.

- Hart, J. A., Reed, R. R., & Bar, F. (1992). The building of the Internet: Implications for the future of broadband networks. *Telecommunications Policy*, 16(8), 666-689.
- Hedgecoe, A., Carvalho, F., Lobmayer, P., & Raka, F. (2006). Research ethics committees in Europe: implementing the directive, respecting diversity. *Journal of medical ethics*, 32(8), 483-486.
- Helbing, D. (2013). Globally networked risks and how to respond. *Nature*, 497(7447), 51.
- Held, V. (2005). The Ethics of Care. In D. Copp (Ed.), *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press.
- Held, V. (2006). *The ethics of care: Personal, political, and global*. Oxford University Press on Demand.
- Henderson, L. (1990). Authoritarianism and the Rule of Law. *Ind. LJ*, 66, 379.
- Herkert, J. R. (2011). Ethical challenges of emerging technologies. In (pp. 35-44).
- Herring, J. (2017). Compassion, ethics of care and legal rights. *International Journal of Law in Context*, 13(2), 158-171.
- Hewison, J., & Haines, A. (2006). Overcoming barriers to recruitment in health research. *bmj*, 333(7562), 300-302.
- Holden, C. (1979). Ethics in social science research.
- Holmes, D. S. (1976). Debriefing after psychological experiments: I. Effectiveness of postdeception dehoaxing. 31(12), 858.
- Howe, K. R., & Moses, M. S. (1999). Chapter 2: Ethics in educational research. *Review of research in education*, 24(1), 21-59.
- Hughes, J., Hunter, D., Sheehan, M., Wilkinson, S., & Wrigley, A. (2010). European textbook on ethics in research., 203.
- Human, D., & Fluss, S. (2003). The World Medical Association's Declaration of Helsinki: Historical and Contemporary Perspectives.
- Hurst, S. A. (2008). Vulnerability in research and health care; describing the elephant in the room? *Bioethics*, 22(4), 191-202.
- Stats, I. W. (2018). Internet World Stats: Usage and Population Statistics.
- Israel, M., & Hay, I. (2006). *Research ethics for social scientists*. Sage.
- Jaimes, A., Sebe, N., & Gatica-Perez, D. (2006). *Human-centered computing: a multimedia perspective*.
- Jasanoff, S. (2007). Making order: Law and science in action. *The Handbook of Science and Technology Studies*. 3rd ed. MIT Press, Cambridge, Massachusetts, 761-786.
- Ji, S., Yang, S., Das, A., Hu, X., & Beyah, R. (2017). *Password correlation: Quantification, evaluation and application*.
- Johnson, D. G., & Powers, T. M. (2005). Computer systems and responsibility: A normative look at technological complexity. *Ethics and information technology*, 7(2), 99-107.
- Jones, A. (1994). The ethics of research in cyberspace. *Internet research*, 4(3), 30-35.
- Jones, B., Ensafi, R., Feamster, N., Paxson, V., & Weaver, N. (2015). *Ethical concerns for censorship measurement*.
- Jones, W. H. S. (2013). *The doctor's oath*. Cambridge University Press.
- Junghans, C., Feder, G., Hemingway, H., Timmis, A., & Jones, M. (2005). Recruiting patients to medical research: double blind randomised trial of "opt-in" versus "opt-out" strategies. *Bmj*, 331(7522), 940.
- Kang, R., Dabbish, L., Fruchter, N., & Kiesler, S. (2015). my data just goes everywhere:" user mental models of the internet and implications for privacy and security.
- Kant, I. (1785). *Groundwork for the Metaphysics of Morals*.

- Karir, M., Huston, G., Michaelson, G., & Bailey, M. (2013). *Understanding IPv6 populations in the wild* International Conference on Passive and Active Network Measurement.
- Karn, P., & Partridge, C. (1987). Improving round-trip time estimates in reliable transport protocols.
- Kaufmann, M., Manjili, A. A., Hildenbrand, S., Kossmann, D., & Tonder, A. (2013). *Time Travel in Column Stores*. Proceedings from ICDE Conference 2013.
- Keller, H. E., & Lee, S. (2003). Ethical issues surrounding human participants research using the Internet. *Ethics & Behavior*, 13(3), 211-219.
- Kelley, P. J. (1991). Proximate Cause in Negligence Law: History, Theory, and the Present Darkness. *Wash. ULQ*, 69, 49.
- Kenneally, E. (2015a). How to throw the race to the bottom: revisiting signals for ethical and legal research using online data. *ACM SIGCAS Computers and Society*, 45(1), 4-10.
- Kenneally, E. (2015b). How to throw the race to the bottom: revisiting signals for ethical and legal research using online data. *ACM SIGCAS Computers and Society*, 45(1), 4-10.
- Kerstein, S. (2019). Treating Persons as Means.
- Khan, M. T., & Kanich, C. (2015). High fidelity, high risk, high reward: using high-fidelity networking data in ethically sound research. *Proceedings of the 2015 ACM SIGCOMM Workshop on Ethics in Networked Systems Research*.
- King, S. A. (1996). Researching Internet communities: Proposed ethical guidelines for the reporting of results. *The Information Society*, 12(2), 119-128.
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1, 1-12.
- Kittay, E. F. (2001). A feminist public ethic of care meets the new communitarian family policy. *Ethics*, 111(3), 523-547.
- Kittay, E. F. (2011). The ethics of care, dependence, and disability. *Ratio Juris*, 24(1), 49-58.
- KNAW. (2016). *Ethical and Legal Aspects of Informatics Research*.
- Knoppers, B. M., & Saginur, M. (2005). The Babel of genetic data terminology. *Nature biotechnology*, 23(8), 925.
- Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N. et al. (2011). *Of passwords and people: measuring the effect of password-composition policies*.
- Kopelman, L. M. (2004). Minimal Risk as an International Ethical Standard in Research. *Medicine and Philosophy*, 29(3), 351-378.
- Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(29).
- Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American psychologist*, 59(2), 105.
- Kreibich, C., Weaver, N., Nechaev, B., & Paxson, V. (2010). Netalyzr: illuminating the edge network.
- Kurose, J. F., & Ross, K. W. (2013). Computer networking: a top-down approach. *Addison Wesley Computing*.
- Lam, V. T., Antonatos, S., Akritidis, P., & Anagnostakis, K. G. (2006). *Puppetnets: misusing web browsers as a distributed attack infrastructure*.
- Lange, L. (2016). Thinking academic freedom. *Arts and Humanities in Higher Education*, 15(2), 177-186.

- LaRossa, R., Bennett, L. A., & Gelles, R. J. (1981). Ethical dilemmas in qualitative family research. *Journal of Marriage and the Family*, 303-313.
- Laudon, K. (1997). Extensions to the theory of markets and privacy: Mechanics of pricing information.
- Lehr, W. (2012). Measuring the Internet: The data challenge.
- Lessig, L. (2006). *Code*. New York: Basic Books.
- Levine, R. J. (1988). *Ethics and Regulation of Clinical Research*. New Haven, CT: Yale University Press.
- Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook. *com. Social networks*, 30(4), 330-342.
- Lewis, M. (2008). New strategies of control: Academic freedom and research ethics boards. *Qualitative Inquiry*, 14(5), 684-699.
- Li, Y., Wang, H., & Sun, K. (2016). *A study of personal information in human-chosen passwords and its security implications*.
- Lloyd, P., & Busby, J. (2003). "Things that went well—No serious injuries or deaths": Ethical reasoning in a normal engineering design process. *Science and Engineering Ethics*, 9(4), 503-516.
- Lowrance, W. (2003). Learning from experience: privacy and the secondary use of data in health research. 8(1), 2-7.
- Luna, F. (2009). Elucidating the concept of vulnerability: Layers not labels. *Feminist Approaches to Bioethics*, 2(1), 121-139.
- Luna, F., & Vanderpoel, S. (2013). Not the usual suspects: addressing layers of vulnerability. *Bioethics*, 27(6), 325-332.
- Ma, J., Yang, W., Luo, M., & Li, N. (2014). *A study of probabilistic password models*.
- MacIntyre, A. (1982). Risk, Harm, and Benefit Assessments as Instruments of Moral Education. *Ethical issues in social science research*, 175-192.
- Malone, D., & Maher, K. (2012). *Investigating the distribution of password choices*. Proceedings from Proceedings of the 21st international conference on World Wide Web.
- Manning, R. (2015). Care, Normativity and the Law.
- Marchant, G. E., Allenby, B. R., & Herkert, J. R. (2011). *The growing gap between emerging technologies and legal-ethical oversight: The pacing problem* (7). Springer Science & Business Media.
- Marco, J. P. (2004). Vulnerability: A needed moral safeguard. *The American Journal of Bioethics*, 4(3), 82-84.
- Margulis, S. T. (1977). Conceptions of privacy: Current status and next steps. *Journal of Social Issues*, 33(3), 5-21.
- Margulis, S. T. (2003). Privacy as a social issue and behavioral concept. *Journal of social issues*, 59(2), 243-261.
- Markham, A., & Buchanan, E. (2012). *Ethical decision-making and internet research: Recommendations from the aoir ethics working committee*.
- Markus, M. L. (1997). The qualitative difference in information systems research and practice. In *Information systems and qualitative research* (pp. 11-27). Springer.
- Matwyshyn, A. M., Cui, A., Keromytis, A. D., & Stolfo, S. J. (2010). Ethics in security vulnerability research. *IEEE Security & Privacy*, 8(2).
- Mazzocchi, F. (2015). Could Big Data be the end of theory in science?: A few remarks on the epistemology of data-driven science. *EMBO reports*, 16(10), 1250-1255.
- McCallister, E., Grance, T., & Scarfone, K. A. (2010). Guide to protecting the confidentiality of personally identifiable information (PII).
- McDonald, A., & Cranor, L. F. (2010). *Beliefs and behaviors: Internet users' understanding of behavioral advertising*.

- McMahon, C. (1991). The paradox of deontology. *Philosophy & Public Affairs*, 350-377.
- McNamara, A., Smith, J., & Murphy-Hill, E. (2018). Does ACM's code of ethics change ethical decision making in software development.
- McNelis, N. (2000). EU communication on the precautionary principle. *J. Int'l Econ. L.*, 3, 545.
- Meara, N. M., & Schmidt, L. D. (1991). The ethics of researching counseling/therapy processes. *Research in counseling*, 237-259.
- Metcalf, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society*, 3(1).
- Milgram, S. (1963). Behavioral study of obedience. *The Journal of abnormal and social psychology*, 67(4).
- Mill, J. S. (1966). On liberty. In *A selection of his works* (pp. 1-147). Springer.
- Mill, J. S. (2015). *On Liberty, Utilitarianism, and other essays*. Oxford University Press, USA.
- Miller, D. (2010). John Stuart Mill.
- Minnis, H. J. (2004). Ethics review in research: Ethics committees are risk averse. *British Medical Journal*, 328(7441), 710-711.
- Minow, M., & Shanley, M. L. (1996). Relational rights and responsibilities: Revisioning the family in liberal political theory and law. *Hypatia*, 11(1), 4-29.
- Molterer, K., Hoyer, P., & Steyaert, C. (2019). A practical ethics of care: Tinkering with different 'goods' in residential nursing homes. *Journal of Business Ethics*, 1-17.
- Montjoye, Y. A. D., Hidalgo, C. A., & Verleysen..., M. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*.
- Moor, J. H. (1985). What is computer ethics. *Metaphilosophy*, 16(4), 266-275.
- Moore, M. S. (2009). *Causation and responsibility: An essay in law, morals, and metaphysics*. Oxford University Press on Demand.
- Moskop, J. C., Marco, C. A., Larkin, G. L., Geiderman, J. M., & Derse, A. R. (2005). From Hippocrates to HIPAA: privacy and confidentiality in emergency medicine—part I: conceptual, moral, and legal foundations. *Annals of emergency medicine*, 45(1), 53-59.
- Mowery, D. C., & Simcoe, T. (2002). Is the Internet a US invention?—an economic and technological history of computer networking. *Research Policy*, 31(8-9), 1369-1387.
- Muller, M. J. (2007). Participatory design: the third space in HCI. In *The human-computer interaction handbook* (pp. 1087-1108). CRC press.
- Narayanan, A., & Felten, E. W. (2014). No silver bullet: De-identification still doesn't work. *White Paper*, 8.
- Narayanan, A., & Shmatikov, V. (2006). How to break anonymity of the netflix prize dataset.
- Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets., 15.
- Narayanan, A., & Shmatikov, V. (2009). De-Anonymizing Social Networks. *IEEE Symposium on Security and Privacy*, 173-187.
- Narayanan, A., & Shmatikov, V. (2010). Myths and fallacies of personally identifiable information. *Communications of the ACM*, 53(6), 3.
- Narayanan, A., & Shmatikov, V. (2005). *Fast dictionary attacks on passwords using time-space tradeoff*.
- Narayanan, A., & Zevenbergen, B. (2015). No encore for encore? ethical questions for web-based censorship measurement. *Ethical Questions for Web-Based Censorship Measurement* (September 24, 2015).

- Commission, N. (1978). The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research-the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.
- Naughton, J. (2016). The evolution of the Internet: from military experiment to General Purpose Technology. *Journal of Cyber Policy*, 1, 5-28.
- Naylor, D., Schomp, K., Varvello, M., Leontiadis, I., Blackburn, J., López, D. R., et al. (2015). Multi-context TLS (mcTLS): Enabling secure in-network functionality in TLS., 199-212.
- Nielsen, L. B. (2010). The need for multi-method approaches in empirical legal research.
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington law review*, 79(1).
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, 140(4), 32-48.
- Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.
- Noddings, N. (1982). *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley: University of CA Press.
- Noddings, N. (2002). *Starting at home: Caring and social policy*. Univ of California Press.
- Nortvedt, P., Hem, M. H., & Skirbekk, H. (2011). The ethics of care: Role obligations and moderate partiality in health care. *Nursing ethics*, 18(2), 192-200.
- NSF. (2011). Other Post Award Requirements and Considerations. *National Science Foundation*.
- Code, N. (1949). The Nuremberg Code. *Trials of war criminals before the Nuremberg military tribunals under control council law*.
- Ogburn, W. F. (1922). *Social change with respect to culture and original nature*. BW Huebsch, Incorporated.
- Ohm, P. (2009). Broken Promises of Privacy: Responding to the Surprising Failure Anonymization. *UCLA Law Review*, 77.
- Öhman, C., & Floridi, L. (2017). The political economy of death in the age of information: a critical approach to the digital afterlife industry. *Minds and Machines*, 27(4), 639-662.
- Öhman, C., & Floridi, L. (2018). An ethical framework for the digital afterlife industry. *Nature Human Behaviour*, 2(5), 318.
- Osuji, P. I. (2018). Relational autonomy in informed consent (RAIC) as an ethics of care approach to the concept of informed consent. *Medicine, Health Care and Philosophy*, 21(1), 101-111.
- Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., & Guston, D. (2013). A framework for responsible innovation. *Responsible innovation: managing the responsible emergence of science and innovation in society*, 31, 27-50.
- Parameswaran, M., & Whinston, A. B. (2007). Social computing: An overview. *Communications of the Association for Information Systems*, 19(1), 37.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Parker, D. B., & Barrett, R. J. (2003). Collective danger and individual risk: cultural perspectives on the hazards of medical research. *Internal medicine journal*, 33(9-10), 463-464.
- Partridge, C., & Allman, M. (2016). Ethical considerations in network measurement papers. *Communications of the ACM*, 59(10), 58-64.
- Partridge, C., & Allman, M. (2016). Ethical considerations in network measurement papers. *Communications of the ACM*, 59(10), 58-64.
- Patterson, M. M. (2010). *Encyclopedia of Research Design*, 333-336.

- Paxson, V., Mahdavi, J., Adams, A., & Mathis, M. (1998). An architecture for large scale Internet measurement. *IEEE Communications Magazine*, 36(8), 48-54.
- Peden, B. F., & Flashinski, D. P. (2004). Virtual research ethics: A content analysis of surveys and experiments online. In E. A. Buchanan (Ed.), *Readings in virtual research ethics: Issues and controversies* (p. 26). Hershey, PA: Information Science Publishing.
- Pettit, P. (1991). *Instituting a research ethic: Chilling and cautionary tales*. Academy of the Social Sciences in Australia.
- Pettit, P. (1997). *Republicanism: a theory of freedom and government*. OUP Oxford.
- Pfaffenberger, B. (1992). Social anthropology of technology. *Annual review of Anthropology*, 21(1), 491-516.
- Pittenger, D. J. (2003). Internet research: An opportunity to revisit classic ethical problems in behavioral research. *Ethics & Behavior*, 13(1), 45-60.
- Plaza, J., & Fischbach, R. Current Issues in Research Ethics: Privacy and Confidentiality [Online Learning].
- Po-Wah, J. T. L. (2002). Is just caring possible? Challenge to bioethics in the new century. In J. T. L. Po-Wah (Ed.), *Cross-cultural perspectives on the (Im) possibility of global bioethics* (pp. 41-58). Springer.
- Poole, E. S., Chetty, M., Grinter, R. E., & Edwards, W. K. (2008). More than meets the eye: transforming the user experience of home network management., 10.
- Poor, N., & Davidson, R. (2016). The Ethics of Using Hacked Data. *Internet Research Ethics for the Social Age*, 277.
- Post, R. C. (2000). Three Concepts of Privacy. *The Georgetown Law Journal*.
- Pound, R. (1910). Law in books and law in action. *Am. L. Rev.*, 44, 12.
- Power, M. (1997). *The Audit Society: Rituals of Verification*. Oxford: OUP.
- Pritchard, I. A. (2001). Searching for "Research Involving Human Subjects": What Is Examined? What Is Exempt? What Is Exasperating. *Ethics & Human Research*, 23(3), 5-13.
- Punch, M. (1994). *Politics and ethics in qualitative research* (2).
- Rabinovich, M., & Allman, M. (2016). *Measuring the Internet* (20(04)). Los Alamitos, CA, USA: IEEE Computer Society.
- Ramcharan, P., & Cutcliffe, J. R. (2001). Judging the ethics of qualitative research: considering the 'ethics as process' model. *Health & social care in the community*, 9(6), 358-366.
- Rangi, S., & Warso, Z. (2015). A report on the legal frameworks that guide or constrain ethical procedures within research within the EU.
- Rawls, J. (1972). 1971. A theory of justice.
- Rawls, J. (2005). *Political liberalism*. Columbia University Press.
- Razaghpanah, A., Vallina-Rodriguez, N., Sundaresan, S., Kreibich, C., Gill, P., Allman, M. et al. (2015). Haystack: In situ mobile traffic analysis in user space., 14.
- Regan, P. M. (1995). *Legislating Privacy Technology, Social Values, and Public Policy*. University of North Carolina Press.
- Reid, E. (1996). Informed consent in the study of on-line communities: A reflection on the effects of computer-mediated social research. *The Information Society*, 12(2), 169-174.
- Reidenberg, J. R. (1997). Lex Informatica: The Formulation of Information Policy Rules through Technology Texas Law Review. *Tex. L. Rev.*, 76(3), 43.
- Resnik, D. B. (2005). Eliminating the daily life risks standard from the definition of minimal risk. *Journal of Medical Ethics*, 31(1), 35-38.
- Rhodes, R. (2005). Rethinking research ethics. *The American Journal of Bioethics*, 5(1), 7-28.

- Rice, T. W. (2008). The historical, ethical, and legal background of human-subjects research. *Respiratory care*, 53(10), 1325-1329.
- Rickham, P. P. (1964). Human experimentation. Code of ethics of the world medical association. Declaration of Helsinki. *British medical journal*, 2(5402), 177.
- Rimm, M. (1994). Marketing pornography on the information superhighway. *Geo LJ*, 83, 1849.
- Rivière, D. (2011). Looking from the outside/in: Re-thinking research ethics review. *Journal of Academic Ethics*, 9(3), 193-204.
- Roberts, C. J. (2014). *Riley v. California*.
- Robinson, F. (1997). Globalizing care: Ethics, feminist theory, and international relations. *Alternatives*, 22(1), 113-133.
- Robinson, F. (2011). The ethics of care: A feminist approach to human security.
- Rogaway, P. (2015). The Moral Character of Cryptographic Work. *IACR Cryptology ePrint Archive*, 1162.
- Rosen, G. (2010). Causation, Counterfactual Dependence and Culpability: Moral Philosophy in Michael Moore's Causation and Responsibility. *Rutgers LJ*, 42, 405.
- Rosenbloom, S. (2007). On Facebook, scholars link up with data., 3.
- Ross, W. D. (2011). *Foundations of ethics*. Read Books Ltd.
- Rudder, C. (2014). We experiment on human beings: OK Trends: Dating Research from OKCupid (Blog). <http://blog.okcupid.com/index.php/we-experiment-on-human-beings>.
- Ruddick, S. (1995). *Maternal thinking: Toward a politics of peace*. Boston, MA: Beacon Press.
- Russell, C. (2002). *Academic freedom*. Routledge.
- Salganik, M. J. (2017). *Bit by Bit: Social Research in the Digital Age*. Princeton University Press.
- Saltzer, J. H., Reed, D. P., & Clark, D. D. (1984). End-to-end arguments in system design. *ACM Transactions on Computer Systems (TOCS)*, 2(4), 277-288.
- Salus, P. H., & Vinton, G. (1995). *Casting the Net: From ARPANET to Internet and Beyond*. Addison-Wesley Longman Publishing Co., Inc.
- Sander-Staudt, M. (2011). Care ethics.
- Schechter, S., & Bravo-Lillo, C. (2014). Ethical-response survey report: Fall 2014.
- Scheffler, S. (1988). *Consequentialism and its Critics*. Oxford University Press on Demand.
- Scheffler, S. (1994). *The rejection of consequentialism: A philosophical investigation of the considerations underlying rival moral conceptions*. Oxford University Press.
- Schrag, Z. M. (2011). The case against ethics review in the social sciences. *Research ethics*, 7(4), 120-131.
- Schramm, W. (1971). Notes on Case Studies of Instructional Media Projects.
- Schroeder, D., & Gefenas, E. (2009). Vulnerability: Too vague and too broad. *Cambridge Quarterly of Healthcare Ethic*, 18(2), 113-121.
- Schrum, L. (1995). Framing the debate: Ethical research in the information age. *Qualitative Inquiry*, 1(3).
- Schwartz, P. M., & Solove, D. J. (2011). The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *NYU Law Quarterly*, 86.
- Sedgwick, S. (2012). *Hegel's critique of Kant*. Oxford University Press on Demand.
- Sevenhuijsen, S. (2003). *Citizenship and the ethics of care: Feminist considerations on justice, morality and politics*. London: Routledge.
- Sharav, V. H. (2003). Children in clinical research: a conflict of moral values. *American journal of bioethics*, 3(1), 12-59.

- Sharkey, S., Jones, R., Smithson, J., Hewis, E., Emmens, T., Ford, T. et al. (2011). Ethical practice in internet research involving vulnerable people: lessons from a self-harm discussion forum study. *Journal of medical ethics*.
- Shaw, S., & Barrett, G. (2006). Research governance: regulating risk and reducing harm. *Journal of the Royal Society of Medicine*, 99(1), 14-19.
- Shuster, E. (1997). Fifty years later: the significance of the Nuremberg Code. *New England Journal of Medicine*, 337(20), 1436-1440.
- Sieber, J. E. (1983). Deception in social research III: The nature and limits of debriefing. *Ethics & Human Research*, 5(3), 1-4.
- Singer, P. A. (1985). Ethics. In. Chicago.
- Singer, P. (1972). Famine, affluence, and morality. *Philosophy & public affairs*, 229-243.
- Singer, P. (2011). *Practical ethics*. Cambridge university press.
- Slote, M. (1999). Caring versus the philosophers. *Philosophy of Education Archive*, 25-35.
- Slote, M. (2007). *The ethics of care and empathy*. Routledge.
- Smart, C. (2000). Children and the transformation of family law. *UNBLJ*, 49, 137.
- Smart, J. J. C., & Williams, B. (1973). *Utilitarianism: For and against*. Cambridge University Press.
- Smith, T. W. (2015). Virtue, Oppression, and Resistance Struggles.
- Snider Jr, D. E., & Stroup, D. F. (1997). Defining research when it comes to public health. *Public Health Reports*, 112(1), 29.
- Solove, D. J. (2006). A Taxonomy of Privacy., 477-560.
- Solove, D. J. (2012). Introduction: Privacy Self-Management and the Consent Dilemma. *Harvard Law Review*, 126, 1880-1903.
- Solove, D. J. (2004). *The digital person: Technology and privacy in the information age*. NYU Press.
- Solove, D. J. (2016). A brief history of information privacy law. *Proskauer on privacy*, PLI.
- Solove, D. J., Rotenberg, M., & Schwartz, P. M. (2006). *Privacy, information, and technology*. Aspen Publishers Online.
- Sperino, S. F. (2013). Discrimination Statutes, the Common Law, and Proximate Cause. *U. Ill. L. Rev.*, 1.
- Spiekermann, S. (2015). *Ethical IT innovation: A value-based system design approach*. Auerbach Publications.
- Spruit, S. L., van de Poel, I., & Doorn, N. (2016). Informed Consent in Asymmetrical Relationships: an Investigation into Relational Factors that Influence Room for Reflection. *Nanoethics*, 10(2), 123-138.
- Stahl, B. C. (2013). Responsible research and innovation: The role of privacy in an emerging framework. *Science and Public Policy*, 40(6), 708-716.
- Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research policy*, 42(9), 1568-1580.
- Strle, G., Benčin, R., Šumič-Riha, J., & Riha, R. (2015). Ethics assessment in different fields: Natural Sciences. *Deliverable 1.1, Annex 2.a*.
- Sveningsson, M. (2001). *Creating a sense of community: Experiences from a Swedish web chat*. Linköpings universitet.
- Tauer, C. A. (2002). Central ethical dilemmas in research involving children. *Policies and Quality Assurance*, 9(3-4), 127-142.
- Thelwall, M. (2001). A web crawler design for data mining. *Journal of Information Science*, 27(5), 319-325.
- Thelwall, M., & Stuart, D. (2006). Web crawling ethics revisited: Cost, privacy, and denial of service. *Journal of the American Society for Information Science and Technology*, 57(13), 1771-1779.

- Thomas, A. (2011). Virtue ethics and an ethics of care: complementary or in conflict. *eidos*, 14), 132-151.
- Thomas, D. R., Pastrana, S., Hutchings, A., Clayton, R., & Beresford, A. (2017). Ethical issues in research using datasets of illicit origin. *Proceedings of IMC '17*.
- Thomas, J. (1995). The Ethics of Carnegie Mellon's "cyber-porn" study.
- Thomas, J. (1996). When cyberresearch goes awry: The ethics of the Rimm" Cyberporn" study. *The Information Society*, 12(2), 189-198.
- Treviño, L. K., Weaver, G. R., Gibson, D. G., & Toffler, B. L. (1999). Managing ethics and legal compliance: What works and what hurts. *California management review*, 41(2), 131-151.
- Tronto, J. C. (1993). *Moral boundaries: A political argument for an ethic of care*. Psychology Press.
- Turow, J., King, J., Hoofnagle, C. J., Bleakley, A., & Hennessy, M. (2009). Americans reject tailored advertising and three activities that enable it. Available at SSRN 1478214.
- Tyler, T. R. (2006). *Why people obey the law*. Princeton University Press.
- Senate, U. S. (1977). Project MKULTRA, The CIA's program of research in behavioral modification. *Joint Hearing Before the US Senate Select Committee on Intelligence and the Subcommittee on Health and Scientific Research of the Committee on Human Resources*. US Government Printing Office, Washington, DC.
- Ur, B., Leon, P. G., Cranor, L. F., Shay, R., & Wang, Y. (2012). *Smart, useful, scary, creepy: perceptions of online behavioral advertising*.
- Ur, B., Segreti, S. M., Bauer, L., Christin, N., Cranor, L. F., Komanduri, S. et al. (2015). *Measuring real-world accuracies and biases in modeling password guessability*.
- Services, U. S. D. O. H. A. H. (2009). Code of Federal Regulations for the Protection of Human Research Subjects (Title 45 CFR 46).
- Vallor, S. (2016). *Technology and the virtues: A Philosophical Guide to a Future Worth Wanting*. Oxford: Oxford University Press.
- van den Hoven, J. (2013). *Options for strengthening responsible research and innovation: report of the Expert Group on the State of Art in Europe on Responsible Research and Innovation*. Brussels: Publications Office of the European Union.
- Van der Burg, W., & Brom, F. (2000). Legislation on ethical issues: Towards an interactive paradigm. *Ethical Theory and Moral Practice*, 3(1), 57-75.
- Van Dijk, N. (2010). Property, privacy and personhood in a world of ambient intelligence. *Ethics and information technology*, 12(1), 57-69.
- Van Roojen, M. (2015). *Metaethics: A contemporary introduction*. Routledge.
- Van Velthoven, B. (2016). A Young Person's Guide to Empirical Legal Research. With Illustrations from the Field of Medical Malpractice. *Law and Method*.
- Varnhagen, C. K., Gushta, M., Daniels, J., Peters, T. C., Parmar, N., Law, D., et al. (2005). How informed is online informed consent. *Ethics & Behavior*, 15(1), 37-48.
- Veatch, R. (1998). The patient as partner: a theory of human experimentation ethics. *Religious Ethics*, 16(1), 190.
- Vellinga, A., Cormican, M., Hanahoe, B., Bennett, K., & Murphy, A. W. (2011). Opt-out as an acceptable method of obtaining consent in medical research: a short report. *BMC medical research methodology*, 11(1), 40.
- Veras, R., Collins, C., & Thorpe, J. (2014). *On Semantic Patterns of Passwords and their Security Impact*.
- Visse, M., Abma, T., & Widdershoven, G. (2015). Practising political care ethics: Can responsive evaluation foster democratic care. *Ethics and Social Welfare*, 9(2), 164-182.
- Walter, J. K., & Ross, L. F. (2014). Relational autonomy: moving beyond the limits of isolated individualism. *Pediatrics*.

- Walther, J. B. (2002). Options for strengthening responsible research and innovation: Report of the Expert Group on the State of Art in Europe on Responsible Research and Innovation. *Ethics and information technology*, 4(3), 205-216.
- Wang, F.-Y., Carley, K. M., Zeng, D., & Mao, W. (2007). Social computing: From social informatics to social intelligence. *IEEE Intelligent systems*, 22(2), 79-83.
- Warwick, D. P. (1982). Types of harm in social research. In *Ethical issues in social science research*. Baltimore: Johns Hopkins University Press.
- Weindling, P. (2001). The origins of informed consent: The international scientific commission on medical war crimes, and the Nuremberg code. *Bulletin of the History of Medicine*, 37-71.
- Weir, M., Aggarwal, S., Collins, M., & Stern, H. (2010). *Testing metrics for password creation policies by attacking large sets of revealed passwords*.
- Weir, M., Aggarwal, S., De Medeiros, B., & Glodek, B. (2009). *Password cracking using probabilistic context-free grammars*.
- Westin, A. F., & Ruebhausen, O. M. (1967). *Privacy and freedom* (1). New York: Atheneum.
- Whitworth, B., & De Moor, A. (2003). Legitimate by design: towards trusted socio-technical systems. *Behaviour & Information Technology*, 22(1), 31-51.
- WHO. (2002). International Ethical Guidelines for Biomedical Research Involving Human Subjects Prepared by the Council for International Organizations of Medical Sciences., 60.
- Wiles, R., & Boddy, J. (2013). Introduction to the special issue: Research ethics in challenging contexts. *Methodological Innovations*, 8(2), 1-5.
- Wiles, R., Crow, G., Heath, S., & Charles, V. (2008). The management of confidentiality and anonymity in social research. *International Journal of Social Research Methodology*, 11(5), 417-428.
- Wilkins, R. G. (1987). Defining the reasonable expectation of privacy: an emerging tripartite analysis. 40, 1077.
- Willinger, W., & Roughan, M. (2013). *Internet topology research redux*.
- Winn, P. (2009). Katz and the Origins of the Reasonable Expectation of Privacy Test. 40, 1.
- Wood, J., Andersson, T., Bachem, A., Best, C., Genova, F., Lopez, D. R. et al. (2010). Riding the wave: How Europe can gain from the rising tide of scientific data. *European Union*, 13.
- Association, W. M. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama*, 310(20), 2191.
- Wright, D. R. (2006). Research ethics and computer science: an unconsummated marriage., 196-201.
- Wright, J., De Souza, T., & Brown, I. (2011). Fine-Grained Censorship Mapping: Information Sources. *Legality and Ethics*.
- Yin, R. K. (2017). *Case study research and applications: Design and methods*. Sage publications.
- Zevenbergen, B. (2016). Networked Systems Ethics. <http://networkedsystemsethics.net/>.
- Zevenbergen, B., Brown, I., Wright, J., & Erdos, D. (2013). Ethical privacy guidelines for mobile connectivity measurements. Available at SSRN 2356824.
- Zhang, B., Liu, R., Massey, D., & Zhang, L. (2005). Collecting the Internet AS-level topology. *ACM SIGCOMM Computer Communication Review*, 35(1), 53-61.
- Zika, E., Schulte In den Bäumen, T., Kaye, J., Brand, A., & Ibarreta, D. (2008). Sample, data use and protection in biobanking in Europe: legal issues.

- Zimbardo, P. G., Haney, C., Banks, W. C., & Jaffe, D. (1972). Stanford prison experiment: A simulation study of the psychology of imprisonment.
- Zimmer, M. (2010). "But the data is already public": on the ethics of research in Facebook. *Ethics Inf Technol*, 12(4), 313-325.
- Zimmer, M., & Kinder-Kurlanda, K. (2017). *Internet research ethics for the social age: New challenges, cases, and contexts*. Bern: Peter Lang International Academic Publishers.
- Zittrain, J. (2008). *The future of the internet--and how to stop it*. New Haven, CT: Yale University Press.