A Framework for Understanding and Applying Ethical Principles in Network and Security Research

Erin Kenneally[1], Michael Bailey[2], and Douglas Maughan[3]

[1] The Cooperative Association for Internet Data Analysis
[2] University of Michigan
[3] US Department of Homeland Security

Abstract. Current information and communications technology poses a variety of ethical challenges for researchers. In this paper, we present an intellectual framework for understanding and applying ethical principles in networking and security research rooted in the guidance suggested by an ongoing Department of Homeland Security working group on ethics. By providing this prototype ethical impact assessment, we seek to encourage community feedback on the working group's nascent efforts and spur researchers to concretely evaluate the ethical impact of their work.

1 Introduction

Innovations in Information and Communications Technology (ICT) have revolutionized how we buy and sell products, how we record, store and playback media, how we communicate with each other, and many other aspects of our lives [4]. Studying the effects of these changes on human welfare, the properties of the enabling technologies themselves, and the ethical implications of the interaction between the two continues to be an active area of study [11, 6]. Expectedly, as the research on impacts of ICT and the enabling technologies become increasingly complex and interconnected, scientists are often posed with moral dilemmas regarding the risks and benefits of such research [7].

One example of a current ICT research (IR) activity that raises novel ethical challenges are efforts to enhance accessibility of computer and network operational data for use in cyber defense research and development. This research acknowledges that the existing lack of practical and reproducible scientific results in ICT research stems in part from a gap between the producers of security-relevant network operations data and researchers who need this data. The PREDICT (Protected Repository for the Defense of Infrastructure against Cyber Threats) initiative of the Department of Homeland Security (DHS) [1] represents an effort to solve this problem. However, the collection and disclosure of networking and security data create a host of dilemmas for those participating in the project and more generally, to all ICT researchers, including: What are user's current perceptions of privacy and confidentiality in network traffic? What are the legal prohibitions to collecting and disclosing network data for research purposes? Is it possible to receive consent by persons implicated in traffic traces? How does one identify a potentially at risk population in a network trace?

Acknowledging the need to resolve these ethical issues not only within its project, but to inform similar debates in other ICT research efforts, DHS hosted a two-day ethics workshop on May 26th-27th, 2009 in Washington, DC[12]. Inspired by the Belmont Report, the 1974 authoritative guide on ethical standards for human subject research[8] in social and behavioral sciences, the workshop brought together ethicists, institutional review boards, researchers, and lawyers to discuss these pressing issues. The primary anticipated outcome from this meeting is a set of ethical guidelines which, though anchored off of the original Belmont framework, reflects the unique questions facing ICT researchers. Subsequently in September and December of 2009, writers working groups met at UC San Diego and Menlo Park, respectively, to advance these guidelines with the intention of publishing them in the first half of 2010.

The goal of this document is to further refine these principles into a workable ethical impact assessment (EIA) that can be used as a framework to help ICT researchers think about the ethical impacts of their work. Unlike work which seeks to answer questions of who should enforce ethical behavior [3, 9] or work that seeks to inform ethical policy debate through the use of case study analysis [7], this work is similar to that of [5, 13] in that we seek to provide specific guidance on how to make ethical research decisions. As the DHS ethics group is a work-in-progress, a secondary goal of this paper is to inform a broader community of this effort and solicit feedback on how to improve the EIA [4].

2 Ethical Impact Assessment (EIA)

In this section, we offer an Ethical Impact Assessment (EIA) framework to more pragmatically assist researchers and evaluators in applying ethical principles in the context of ICT research. This EIA is an incipient prototype, modeled after the more established privacy risk management framework, the PIA (Privacy Impact Assessment) [14]. As such, the EIA offers non-exhaustive, yet directed, questions to guide compliance with the ethics principles that were put forth at the DHS ethics workshop. These ten principles fall into two categories: guidance on human subjects protection and guidance on professional ethics.

2.1 Human Subject Protections

What do the principles of Respect for Persons, Beneficence, and Justice means to ICT research stakeholders? Because these ethical mandates originated within the context of Human Subjects protection research [8], they have been evaluated and appropriately modified and clarified for ICT network and security context.

Respect for Persons In the context of Human Subjects protection work, respect for persons encompasses at least two components: first, that individuals should be treated as autonomous agents, and second, that persons with diminished autonomy are entitled to protection [8]. These are often applied through the construct of *informed consent*, which in the context of ICT networking and

[4] The specific interpretations expressed in this paper are the authors and don't necessarily reflect that of other individual working group participants

security research, raises questions of identification, the appropriate level of disclosure of research methodology, comprehension by subjects via network modalities, and voluntariness. Resolving these questions can be vexing if not impracticable in network contexts, raising debate about whether these are suitable means to achieve informed consent, or even whether this construct itself is an appropriate mechanism to realize respect for persons.

1. In the cyber security context, respect for persons should include both individuals and society, and should consider organizations. Ethical challenges posed by privacy concerns can be vexing for ICT research because the underlying concept of identity in relation to network data artifacts is disjointed in both law and social convention. Unlike well-entrenched identifiers such as name or biometric markers, blanket characterizations of IPA or URLs as personally-identifying (or not) are misguided because they alone do not capture the range of privacy risks associated with network traffic which are referential and context-dependent. Furthermore, it may be difficult or impracticable to identify potentially at risk populations in a network trace, such as with juvenile subjects who may warrant greater protections, not to mention the ensuing challenges to obtaining valid consent. Question(s): Consider how data and computer systems may be tightly coupled with the entities to be respected. Can the IP address or URL be relatively easily linked to an identifiable person? Does the IP address map to an automated device, distinguish a human-operated host, or identify a home computer? Does the content of the collected data concern the substance, purport or meaning of a communication from an identifiable person? Does the data reveal behavioral information that could identify an individual? Researchers should be mindful that individuals' dignity, rights, and obligations are increasingly integrated with the data and IT systems within which they communicate, transact, and in general represent themselves in a cyber context.

2. Consent to use data and information systems for a specific purpose in research should be obtained. The challenging aspect of this precept is that *in vivo* Internet research may involve situations where individual consent is impracticable because it would be legally unwarranted or strategically or economically infeasible to identify persons implicated within network and security research data; or, failure to obtain consent would have no adverse impact on an identifiable person's rights and welfare. Since consent often presumes the existence of an underlying legal right, ambiguity over ownership and control of network traffic– e.g., is it public or subject to an expectation of privacy– may complicate consent obligations. Question(s): If the research involves identifiable individuals, have the individuals implicated in the network and security data consented to involvement? Can the individual decline participation in the research or in the uses of collected data? If the purpose of the data use has changed or is expanded, has renewed consent been obtained? If consent is impossible or directly impedes research goals, consider the risk-utility assessment guidance under Beneficence.

Beneficence The Belmont Report [8] specifies two general rules under the obligation of beneficence: "(1) do not harm and (2) maximize possible benefits

and minimize possible harms." Thus beneficence is applied as a *risk-benefit assessment*. The following EIA questions are intended to elicit what is meant by benefits and harm in the context of ICT networking and security research.

3. Researchers should systematically assess both risks and benefits of the research on privacy, civil rights, and the well-being of persons. Laws are enacted to secure the rights and well-being of individuals and they offer one systematic approach for evaluation. However, risk-benefit determinations can be challenging given gaps and grey areas in privacy and civil rights laws related to liability for actions undertaken in the interests of security research. Other enforcement mechanisms and systematic approaches have been challenged as inappropriate or incomplete including IRBs and professional codes [3, 9]. This lack of concrete guidance, however, does not assuage the responsibility to perform more than a piecemeal or perfunctory ethical analysis of a study's impact. Question(s): What are the effects of network and security research on all the stakeholders: researchers, human subjects, and society (by way of how it may assist attackers)? In what circumstances will the benefits of the IR clearly outweigh any harmful impact on the stakeholders? Will the research result in no greater harm than what would have occurred in its absence? What checks and balances are in place to prevent both new harms and/or repeated historical abuses, including: violating the law and privacy interests; targeting and disrupting certain groups (based on politics, race, sex, etc.); chilling First Amendment rights (e.g., free speech, freedom of association); harming individuals (e.g., physical, financial, legal, reputational, mental); impairing data quality and integrity (e.g., distorting data that informs government policy or public perception); creating a high cost-to-effectiveness study; introducing surveillance harms (e.g., identity theft, disclosure of embarrassing information, government persecution, chilling or foregoing certain activities, introducing costs or altering behavior related to counter-surveillance); and, expanding network surveillance and perpetuating secrecy.

4. Research should be designed and conducted to maximize probable benefits and minimize harms to persons and organizations. Prominent application challenges here include the scale at which risk and benefits can occur, the ability to attribute research data and results to specific individuals and/or organizations, the increasing availability of data that are beyond the knowledge or control of the researcher (thereby challenging the effectiveness of disclosure controls), and the increasingly intertwined nature of the physical and virtual worlds. This principle seemingly imagines ex ante benefit maximization and risk minimization for research whose value may be conjectural and delayed and whose risk posture and mitigation may be speculative. As such, the following questions help to align expectations and capabilities into practical focus for researchers. Question(s): Does the research impact the integrity, confidentiality, and availability of information systems, including originating and transiting systems? Does the research design include controls to minimize harms and maximize benefits such as using test environments, anonymization techniques or other disclosure controls that limit the exposure of personal data? For example: What are possible unintended conse-

quences of the IR? Are there exigent circumstances that should be factored into the evaluation of harm? Are there privacy-based harms from IR? What is the nature of the information collected by IR? What is the purpose for collecting the data? What is the intended use of information collected by IR? Will the research be disseminated to third parties and used consistent with its original purpose? What are the administrative and technical controls? In assessing the risk of re-identification, consider variables such as: triggers set by law or policy guidelines (e.g., highly probable, readily ascertainable, likely); the quantity of data that would be available; the threat perspective (e.g, a subjective person associated with the data, an objective member of the public, a motivated intruder); and, the level of time, effort and resources needed to re-identify a person.

5. If research reveals or causes risk/harm to a person, including systems and data, the person should be notified. ICT do not often require human interaction or human notification to cause harm or do good. As such, we have a special obligation to inform, where reasonable to do so, those individuals or organizations whose resources and welfare are affected by the phenomena we are measuring. Question(s): When notification of persons is not possible or appropriate, harm should be mitigated by notifying other appropriate parties.

6. Researchers should consider the full spectrum of risks of harm to persons and information systems, including reputational, emotional, financial, and physical harms. Significant here is our normative social immaturity regarding qualitative and quantitative assessment of damages and harms in the electronic realm, as opposed to the well-established and socially-embedded understanding of cause and effect harms resulting from physical interactions with human subjects. Question(s): What categories of activity have especially strong reasons for IR involving human subjects? Could the IR actually make the targeted problem (e.g., security) worse or undermine the research goal(s)?

Justice In the context of human subjects research protection, Justice addresses fairness in determining who ought to receive the benefits of research and bear its burdens [8]. It is thus applied through the construct of *selection of subjects.* While most of these questions do not vary significantly for ICT networking and security research, their application, nonetheless, introduces previously addressed challenges related to identification of persons from referential network data, as well as difficulties in projecting results of research activities involving tightly coupled network systems.

7. The benefits and burdens of research should be shared fairly between research target subjects and beneficiaries of the research results. Question(s): Does the IR raise fairness and discrimination concerns? Will the IR undermine cooperation from the community whose cooperation/participation is needed/targeted?

8. The selection of research subjects should be equitable, except when biased selection may be beneficial. Question(s): To what extent does the IR violate legal

and ethical principles of equality? How can research design be altered to decrease the inequality or mitigate its effects?

2.2 Professional Ethical Guidance

Professional organizations such as IEEE and ACM offer professional codes of ethics for their members [10, 2] and the primary difference between these codes and codes for protection of human subjects is that while these codes recognize an imperative for their member to do good, these codes focus on workplace and employment-related ethical situations rather than on the experimental subjects.

9. Research activities should not violate laws, operator agreements, contractual obligations, or other restrictions agreed to by private arrangements. This consideration ensures that researchers engage in legal due diligence for activities that occur outside of a closed, self-contained research setting and which are subject to laws or policies intended to protect individual and organizational rights. This provision may prove challenging in light of the uncertain application or interpretation of certain laws and regulations in the context of ICT research activities, including the heightened risk of unanticipated consequences or discoveries involved in *in vivo* ICT research. Question(s): If the IR is in conflict with law or policy, is there an exception or valid agreement otherwise permitting such research? Would the IR violate other countries' laws? If government is involved, will there be international and bilateral diplomatic ramifications? Should the IR methodology be modified or abandoned wholesale because of legal and other concerns?

10. Where possible, researchers should adhere to internationally accepted best practices and standards in conducting research and assessing risk. Similar to legal risk assessment involving domestic laws, international risk assessment may be even less clear given the discrepancies between nation-states on cyberlaws and rights. Again, the standard against which research should be measured is that of a reasonable researcher, and not a strict liability. Adherence to international standards or guidelines can often move researchers beyond ethical risks when laws are unclear or unsettled.

3 Conclusion

Increasingly, networking and security researchers are engaging in work that challenges our existing ethical frameworks. If we are to continue to occupy a moral high ground in which we claim the benefits of our work as necessary and the risks of our work minimal, we need to more explicitly justify this reasoning to other researchers and society as a whole. In this paper, we discuss an evolving Ethical Impact Assessment, based on the collaborative efforts of a DHS ethics working group, that seeks to define a set of imperatives for networking and security research. Used as an intellectual framework, it offers the promise of guiding researchers to ask the appropriate set of questions about their work and reason effectively about its ethical impact. As a living document, the authors and working group members actively solicit community feedback on this effort.

References

1. Protected repository for the defense of infrastructure against cyber threats (PREDICT). http://www.predict.org.
2. ACM Council. Code of Ethics and Professional Conduct, October 1992. http://www.acm.org/about/code-of-ethics.
3. Mark Allman. What ought a program committee to do? In *WOWCS'08: Proceedings of the USENIX Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems*, pages 1–5, 2008.
4. Sara Baase. *A Gift of Fire: Social, Legal, and Ethical Issues in Computing*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2002.
5. Aaron J. Burstein. Conducting cybersecurity research legally and ethically. In *LEET'08: Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*, pages 1–8, 2008.
6. Terrell Ward Bynum and Simon Rogerson. *Computer Ethics and Professional Responsibility: Introductory Text and Readings*. Blackwell Publishers, Inc., Cambridge, MA, USA, 2003.
7. David Dittrich, Michael D. Bailey, and Sven Dietrich. Towards community standards for ethical behavior in computer security research. Technical Report 2009-01, Stevens Institute of Technology, Hoboken, NJ, USA, April 2009.
8. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The belmont report - ethical principles and guidelines for the protection of human subjects of research. http://ohsr.od.nih.gov/guidelines/belmont.html.
9. Simson L. Garfinkel. IRBs and security research: Myths, facts and mission creep. In *Proceedings of UPSEC '08 (Usability, Psychology and Security)*, April 2008.
10. IEEE Board of Directors. IEEE Code of Ethics, February 2006. http://www.ieee.org/portal/pages/iportals/aboutus/ethics/code.html.
11. Deborah G. Johnson and Keith W. Miller, editors. *Computers Ethics*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2009.
12. Erin Kenneally. What's belmont got to do with it? http://blog.caida.org/best_available_data/2009/06/12/whatelmont-got-to-do-with-it/.
13. Erin Kenneally and K. Claffy. An internet sharing framework for balancing privacy and utility. In *Engaging Data: First International Forum on the Application and Management of Personal Electronic Information*. MIT, IEEE, October 2009.
14. DHS Privacy Office. DHS official privacy impact assessment guidance. http://www.dhs.gov/files/publications/gc_1209396374339.shtm.