Bridging the Distance: Removing the Technology Buffer and Seeking Consistent Ethical Analysis in Computer Security Research.

By Katherine J. Carpenter*
and David Dittrich ‡

* corresponding author; University of Denver, Denver, CO.
‡ University of Washington, Seattle, WA.

Bridging the Distance: Removing the Technology Buffer and Seeking Consistent Ethical Analysis in Computer Security Research.

Abstract:

Computer science (CS) research in general, and computer security research in particular, often involve manipulating complex systems that have the potential to affect large numbers of individuals. Due to the complicated nature of these interconnected systems, those at risk of harm are not only the end-user owners of computers, but also service providers and other enterprises that are intermediaries in delivery of technological services. This results in an intricate mixture of stakeholders creating confusion determining what truly is "human subjects research" and thus what requires oversight by institutional research ethics boards.

In the United States, when researchers "obtain[] (1) data through intervention or interaction with the individual, or (2) identifiable private information" they are conducting human subjects research. (45 CFR 46. 102(f)). Research is generally designed to be as transparent as possible at least for the researchers and their funding agencies. However, the nature of computer security research diminishes this transparency and creates a distance between the researchers and potentially impacted parties. When computer science researchers believe they are not interacting with the humans their research impacts, they may not consider their research activities to be "human subjects research." There are not consistent ethical standards for considering and measuring the adverse effects of computer science research on human subjects or society.

Broad groups that are affected by computer research include society and criminals/attackers. These categories can be narrowed to: researchers and their programmers; vendors who use the internet to sell products or internet service providers (ISPs); the programmers for vendors; clients and customers of websites, online stores and ISPs; and criminals who exploit internet-based services and/or the data that they are able to discover through technological vulnerabilities.

Due to a history of abuse in research much of the focus of research review focuses on informed consent of participants. In CS research this is virtually impossible. Anyone with an internet connection can be a participant. It would be impractical and unfeasible to obtain informed consent from all individuals and likewise impossible to inform them of the risks. We propose an ideological shift from focusing on informed consent to potential human harms that each CS research project may present. Since it is difficult to determine whom CS research will affect, considering the potential for harm in the research design process and creating strategies to minimize that potential harm is our focus.

Bridging the Distance: Removing the Technology Buffer and Seeking Consistent Ethical Analysis in Computer Security Research.

I. Introduction:

Computer science research is taking place throughout the United States and the world. When average people think about this research, they rarely consider the range of lives touched by access to technology or the ethical concerns raised and still being explored. In the United States Institutional Review Boards (IRB) review most research projects, especially if it is federally funded. Other countries have similar entities called Research Ethics Boards (REBs). Many computer science researchers do not perceive their research activities as having human subjects. Academic, military, government, and private institutions sponsor computer science research. We focus our discussion on the academic setting, but our points have implications for the military, government, and in the private arena.

Most social or medical research on humans is characterized by the degree of interaction between researchers and research participants, also known as "human subjects of research." Research oversight usually entails a discussion of the harms that could come to human subjects from the research and monitoring harms once the research begins. In computer security research, the potential participants are difficult to anticipate and harm to computer end-users is not immediately apparent even though the potential for harm is often fairly high. When researchers test computer systems for problems, and/or combat or mimic viruses, botnets/other malicious actions, the computer systems can be compromised as well as personal computers or data, causing harm.

Computer science research in general, and computer security research in particular, often involve manipulating complex systems that have the potential to affect

large numbers of individuals. Due to the complicated nature of these interconnected systems, those at risk of harm include computer owners, service providers, and other intermediaries that deliver technological services.[i] This results in an intricate mixture of potential participants creating confusion regarding the status of computer security research. Does it qualify as "human subjects research" and require oversight by review boards?

We propose that this type of research needs review, but a shift in research analysis and oversight in this arena is required. Review boards should transition from an informed consent driven review to a risk analysis review that addresses potential harms stemming from research in which a researcher does not directly interact with the at-risk individuals. Computer science researchers may be reluctant to bring forward research that could negatively impact individuals indirectly. Computer science research is important and should be encouraged. Regulatory reform is needed to provide guidance to researchers and review boards.

II. Does CS Research Need Review?

Before deciding how to review research we must first establish that CS researchers undertake research endeavors that actually require review. For example, computer security researchers may look for vulnerabilities in software so they can repair them. They may mimic viruses and botnets[1] in order to understand them, eliminate them, and identify ways to clean infected machines. Research may include collecting personal data from compromised machines.

"Research means a systematic investigation, including research development,

[1] Individual infected computers (known as "bots", short for "robot") form distributed malware systems used by criminals (known collectively as a "botnet," short for "robot network").

testing and evaluation, designed to develop or contribute to generalizable knowledge."[ii] Computer science researchers often experiment with goals to improve the technology environment and to break new ground in a rapidly advancing field of technology

In the United States, when researchers obtain "(1) data through intervention or interaction with the individual, or (2) identifiable private information" they are conducting human subjects research.[iii] The federal regulation, 45 CFR 46 clarifies: "[i]intervention includes both physical procedures […] and manipulations of the subject or the subject's environment that are performed for research purposes." Some commenters have suggested there are definitional problems with some key terms in federal regulations.[iv] We believe the terms "intervention" and "subject's environment" may equally deserve reconsideration and redefinition when it comes to determining when greater than minimal risk of harm to humans within CS research.

There is a great deal of computer security research taking place at universities across the United States and the majority of it is not reviewed by institutional review boards that oversee research with human participants. Even when research is reviewed, the review may not be effective if neither the researchers nor the reviewers have the expertise to consider the indirect impact on the end users of technology as "participation" in research.[v] Researchers may not consider the potential harm that the technology could cause to end users, and many review board members are unaware of the potential risks technology users face.

One pillar of ethical research is the informed consent of participants.[vi] In computer security research, the range of participants could be anyone with an Internet connection. With such a broad base of potentially impacted individuals, informed

consent may not be possible.

The Belmont Report[2] accompanies federal regulations as a guideline for ethical research. The report identifies three ethical principles worthy of consideration in "all human subjects research: respect for persons, beneficence, and justice."[vii] We only deal with the first of these, respect for persons, because it encompasses informed consent. Respect for persons is intended to ensure that research participation is voluntary and vulnerable populations are protected. "In computer security research, unless researchers are collecting data, they do not ask for consent because they do not directly interact with anyone."[viii]

We would like to bring attention and encourage analysis of the harms that can result from this type of research. We recognize that computer security research is important and we urge researchers and review boards to consider the resultant harms that could occur to anyone connected to the Internet because of their research.

Many understand harm as physical or mental damage[ix] or injury. In the context of computer security research harm to software is not physical and there are few ways to physically damage a computer over the Internet. Some potential harm mentioned at the example in the beginning of this section includes compromising private, identifiable information. Identifiable information can be names, identification numbers (including social security or medical record numbers), birthdates, phone numbers, and photographs. The Health Insurance Portability and Accountability Act (HIPAA) has named 18 identifiers as protected health information including the aforementioned identifiers and

[2] The National Commission on the Protection of Human Subjects of Biomedical and Behavioral Research published "Ethical Principles and Guidelines for the Protection of Human Subjects of Research," in 1978. Their report is commonly called The Belmont Report after the Belmont Conference Center where the Commission met to discuss and draft the report.

universal resource locators (URLs), Internet Protocol (IP) addresses. HIPAA is the standard because health research is the most heavily regulated type of research in the US.

Other possible harms stem from the temporary disruption of computers or networks. If a computer is unresponsive or the network is not available when it is expected to be running, "harm" can be severe-- negatively affecting reputation, loss of revenue, disruption of government services or processes like an election. The more humans rely on information technology as a foundation of their lives, the potential for human harm from research, whether direct or indirect, increases.

1. Why are these protections insufficient?

If technology is the subject of research, the humans who use the technology are not considered “research subjects” but they may be subject to harm directly caused by the research because they use technology. This is true whether the technology is a computer virus or an embedded medical device. Both pieces of technology, viewed in a vacuum, appear to have no impact on humans. In reality humans interact with both types of technology and are endangered if that technology is disrupted, either on purpose or by accident. Current protections in place for research are insufficient because many IRB members and researchers do not consider end results in research study design. Federal research guidelines do not account for research that harms individuals without directly interacting with them.

2. IRBs believe they deal with distance/indirection but they really don’t

Researchers, administrators and IRB members focus on efficient mechanisms for review. A report from a 2003 Illinois conference recommended, “focusing on those areas of research that pose the greatest risk, such as bio-medical research, while removing or

reducing scrutiny of many fields within the social sciences and humanities that pose minimal risk."[x] This recommendation does not consider whether researchers or IRBs can adequately assess the level of risk that research utilizing ICT poses to humans on the Internet, whether those humans are the subjects of research or are simply reliant upon studied ICT resources. We do not advocate increasing the scope of IRB review simply to cover new fields, but to accurately identify and review research protocols that pose greater than minimal risk to humans.

Most computer security research is not reviewed as human subjects research, [xi] and if it is, it reviewed as "minimal risk" or via expedited review procedures that do not require full committee evaluation. Review boards often do not understand the implications of the research or the potential for human harm. IRBs have ample access to medical experts, both on boards and as voluntary consultants, but they often do not have the technical expertise available to evaluate CS or CompSec research protocols. Other commenters have pointed out potential mechanisms for altering the balance of biomedical vs. ICT technical expertise by changing the incentive structure for CS researchers so they either sit on IRB committees, or at least are known sources of technical expertise for reviewing CS research.[xii] This results in a gap of understanding of the actual risks presented by the research protocols. The research review process in the United States focuses on the potential harm to *human subjects* of research and the research subjects in most CS research studies are technological tools (like computers).

The distance between researchers and harmed humans mentioned perpetuates the lack of research review. Our goal is to bridge the distance between researchers and (potentially) harmed human users of technology by analyzing the potential for harm in

varying technological research and building protections into the research design.

II. Understanding Potential Harms in Computer Security Research

A. Computer Science and Computer Security Research

Computer Science (CS) is a branch of science that deals with the theory of computation, the design of computer hardware and software, and the use of computing technology in a variety of fields. It is becoming difficult to find any business, government, personal, or research activity that does not utilize computer systems. We focus on the potential harm that results from a connection to the Internet. Information collection, use, and disclosure (i.e., confidentiality concerns) merely begins to address the risks posed by the presence of ICT in research. Risks include breaching confidentiality and compromising the availability and integrity of information and information systems. CS research can be passive—observational—or active. The risk of potential harm is greater in active research. We describe examples of both types below.

Researchers in the sub-discipline of computer security (CompSec) research attempt to measure and control the tactics, strategies, design, operation and application of software or computing systems to protect them from malicious or criminal manipulation. CompSec researchers are interested in discovering the vulnerabilities in certain types of software and other technology or in learning how attacks take place and how to prevent them. Malicious software, known as "malware," is the subject of research aimed at reducing the amount of criminal activity that occurs on the internet. Malware can take many forms, but the one we will discuss here involves mass infection of third-party computers by criminals to create malicious botnets.

With subcomponents of software engineering, network engineering and systems

engineering, CS is implicitly an engineering discipline. Engineering could be considered a form of experimentation on other humans. Martin and Schinzinger suggest:

> ... engineering should be viewed as an experimental process. It is not, of course, an experiment conducted solely in a laboratory under controlled conditions. Rather, it is an experiment on a social scale involving human subjects.[xiii]

Engineering may be considered experimentation, but most engineering falls outside the purview of Federal research regulations, including CS research. Academic researchers who test technology fail to recognize the way their research constitutes experimenting on humans.

B. Examples of Computer Security Research

Many types of research fall under the umbrella of "computer security" research. We address four ethically challenging forms.

1. Vulnerability Research

Vulnerability research identifies technical weaknesses in systems that could allow an adversary to bypass security mechanisms and assume control or elevate privileges within a protected system. This research is important because software vulnerabilities are exploited to gain control of computers through an infection mechanism that creates bots and botnets, or to find avenues that could be used to disrupt or disable building control systems, embedded medical devices, automobile control systems, or electronic voting systems.

There is a long-standing controversy surrounding vulnerability disclosure, with two extreme viewpoints: full disclosure of vulnerabilities upon discovery, and non-

disclosure. Researchers may publicly disclose information before contacting the producers of vulnerable systems. This early disclosure is based on the researcher's desire to get maximum press exposure, to avoid legal pressure from the affected vendor. Researchers may believe affected vendors will not act unless the public is fully informed about the ways the vulnerability could be exploited to cause harm, and disclosure engages market forces that compel the vendor to respond.

Formalized vulnerability disclosure guidelines go back to the early 2000s. Carnegie Mellon University's CERT/CC[3] was one of the first programs to publish vulnerability disclosure guidelines. Other guidelines came from the Organization for Internet Safety, the National Infrastructure Advisory Council, and computer security firms and software vendors over the years.[xiv]

A high profile example of a successful coordinated disclosure of fundamental vulnerabilities in the Domain Name System (DNS) services[xv] relied upon by all Internet users occurred in 2008. Hundreds of vendors were involved over several months to ensure that the majority of them were prepared to release patches fixing the problem on or about the same time as a coordinated public disclosure occurred. A blog post in 2009[xvi] publicly acknowledged Microsoft's coordinated vulnerability disclosure activities and processes. A full statement of the policy indicating how Microsoft could assist security researchers with coordinated disclosure was released in 2011. [xvii]

2. Botnet Takedown Research

Researchers at universities, security software vendors globally, and private individuals actively engage in botnet research and takedown activities. These actions

[3] Originally "Computer Emergency Response Team," the name changed in the mid-2000s to "CERT/CC [TM]" because the coordinating center (CC) was the brand they preferred to maintain. http://www.cert.org.

benefit society by taking control of stolen computing assets out of the hands of miscreants who may be causing significant financial harm. Effectively botnet takedowns become "good guys" fighting with "bad guys" to control stolen computing assets belonging to innocent victims who may be unaware their computers are compromised. These takedowns pose additional risks to computers infected with malware above and beyond monitoring risks in passive network traffic monitoring studies.

The Department of Commerce and National Institute of Standards and Technology (NIST) published a notice in the Federal Register promoting "Models to Advance Voluntary Corporate Notification to Consumers Regarding the Illicit Use of Computer Equipment by Botnets and Related Malware." This notices cites researchers who "suggest an average of about 4 million new botnet infections occur every month." They mention other efforts to involve ISPs in helping clean up infected computers, such as the Internet Engineering Task Force (IETF) draft, "Recommendation for the Remediation of Bots in ISP Networks." [xviii] The notice indicates several international governmental programs dedicated to prevent botnet infections from spreading. Botnet infections are a public threat and the public is served by preventing them.

The Commerce Notice points out that, "Internet Service Providers (ISPs) [have] contact information for the end-user and a pre-existing relationship" with them that could facilitate communicating with end users. Utilizing the ISP relationship may not be a suitable mechanism for obtaining informed consent, as the number of users involved is still large enough as to make it impractical. It would be a costly mechanism to obtain informed consent from all end users. However, we recognize that ISPs *could* act as proxies of consent for their clients.[xix]

The public wants to feel safe:

Public pressure may encourage researchers to clean up infected computers without involving end users. Recently, Microsoft obtained an ex parte temporary restraining order allowing the company to disable thousands of malicious domain names and sinkhole the Kelihos botnet with Kaspersky Labs researchers.[xx] An unscientific poll by Kaspersky[xxi] asked, "How should Kaspersky proceed with the Hlux/Kelihos Botnet?" 78 percent of respondents [4] believed Kaspersky should, "Push a cleanup tool that removes the infections" as opposed to doing nothing or working with ISPs to notify infected victims. Pressure from a frustrated general public, combined with a lack of generally accepted guidelines for researchers to consider such risky actions, may push researchers to take high risks that could potentially damage computers or data in their effort to stem what the public views as widespread harm by criminals.

3. Electronic Voting Research

The electronic vote tabulation systems used in many states employ computer hardware that is very similar to the personal computers in our homes, running similar popular commercial operating systems on which custom software that presents options to a voter. These computers record accumulated votes and deliver them to a central tabulation system that provides the election results using network connections. Several research groups have examined, both the hardware and software of these systems to prevent elections occurring on flawed equipment.

Elections can be, and have been, manipulated by those who wish to serve their political interests. Flaws that can be remotely exploited like manipulating votes, altering

[4] 1159 individual responses were recorded in this unscientific, blog-based poll as of

counts, and bypassing audit checks, pose a serious threat to the integrity of the voting system.

4. Cyber-Physical Systems Research

The term "cyber-physical systems" (CPS) broadly applies to any type of computing device that can be controlled remotely through a computer control system and has direct physical effects on objects in the physical world. This can include process control systems in factories, embedded medical devices that control heart rhythm or delivery of drugs like insulin, and brake systems in automobiles. Research on these types of devices would almost never undergo IRB review because humans are not involved when researchers simply study the device's function outside of their intended use. A vulnerability discovered in a cyber-physical system that can be exploited remotely, using low-cost equipment, could allow a malicious actor to control these systems and literally kill someone.

1. Refining our idea of distance

The end users of tested technology are people and if/when harm occurs to the computer or system because of the research, the human user is often affected negatively. This distance between researcher and affected individual indicates that a paradigm shift is necessary in the research arena. We must transition our idea of research protection from "human subjects research" to "human harming research."

When ICT is present in a research setting, the risk of harm may be much broader than in direct intervention research, where the research subject is the primary party at risk. This is illustrated in Figure 1.

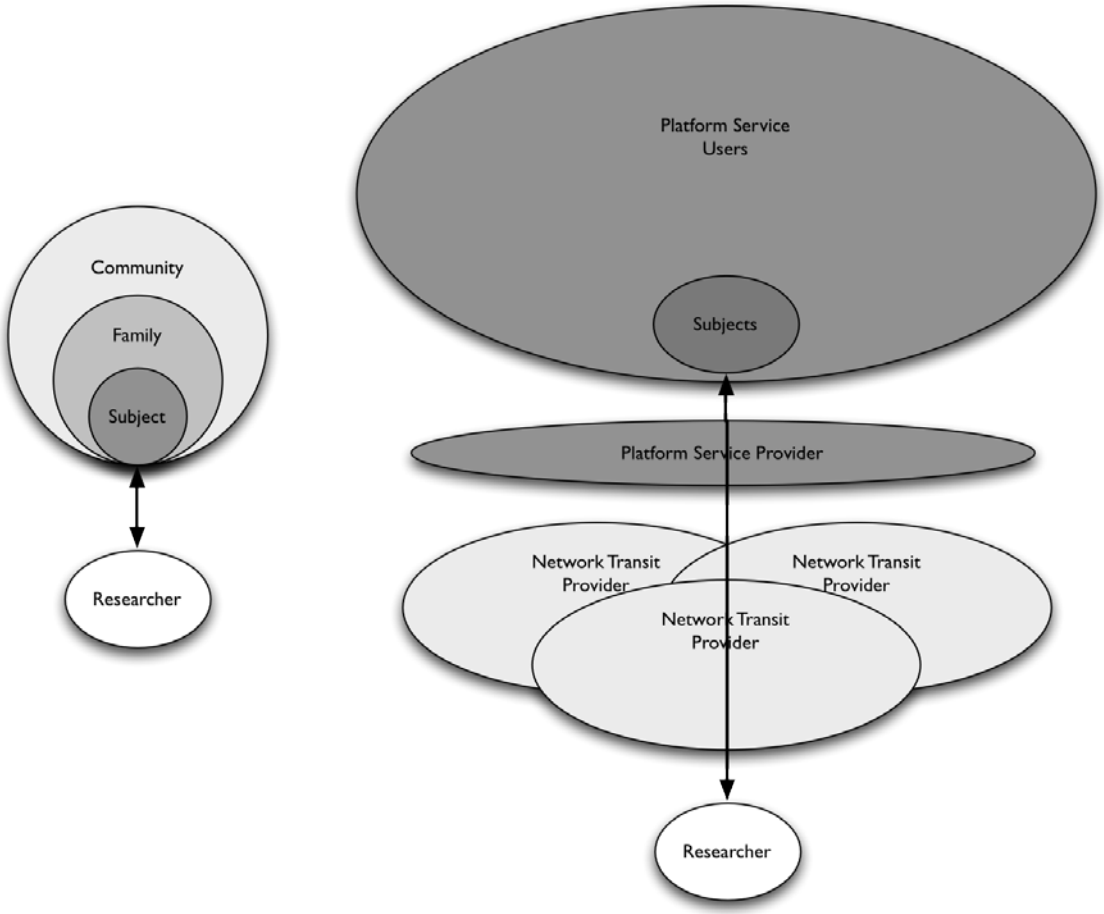

Figure 1.

The arrow in Figure 1 illustrates the relationships between researcher and "human subject." The amount of gray indicates relative level of potential risk (lighter meaning less risk, darker meaning more risk) to stakeholders.

The left side of Figure 1 shows the risk relationships in a typical biomedical or behavioral research study. In this example, the researcher interacts directly with the subject and potential harm extends indirectly to other parties and decreases in severity (although not necessarily in total number of those impacted) as you move away from the subject through that individual's relationships.

The right side of Figure 1 shows the risk relationships in a study that uses ICT. In

researcher- subject interaction, intermediary providers of ICT services and the subjects are a subset of the entire user base of a given platform provider (e.g., a social network site). Notice how the degree of risk is inverted from the previous example. There may be little or no risk to transit providers, but the platform provider and/or other users in the social network of the subjects being studied, may at nearly the same risk level as the research subjects.

2. Many CS researchers don't believe their research is human subjects research

CS researchers may not consider their activities to be "human subjects research," resulting in a dearth of regular and consistent assessment of adverse effects to systems or individuals. The CS community does not have consistent ethical standards for considering and measuring the adverse effects of research on stakeholders.[xxii]

People interact with Technology (and vice-versa):

CS and CompSec research is not limited to direct or manual interaction with human end users, it engages them indirectly. *Intervention* is not limited to physical procedures. Intervention can be "manipulations of the [subject's] environment that are performed for research purposes," including manipulation of their computing devices or automated appliances at home.

Interactions with humans in virtual environments using ICT impacts exponentially more individuals than a direct intervention would. Consider a behavioral study performed within an online virtual world environment observing avatar interactions. Some may argue an avatar is a graphical construct representing a human, but is not a "living individual" who can "interact" with the researcher and thus is never subject to ethical review. The same argument is used for researching malicious software,

embedded medical devices, or a process control device in a dam preventing water from flooding cities. Despite the fact that humans are not the *direct subjects* of the research, the research may involve greater than minimal risk to humans.

Computer science and computer security research do not seem like human subjects research because technology appears to act as a buffer between researchers and individuals. This buffer seems present even though research may directly impact individual users of a system or compromise their computer. Research is generally designed to be as transparent as possible at least for the researchers and their funding agencies. The nature of ICT research diminishes transparency and creates a distance between the researchers and potentially impacted parties (be they direct participants, or indirectly involved "research subjects," because ICT itself is the subject of research).

The following diagram compares three research contexts to illustrate this concept.

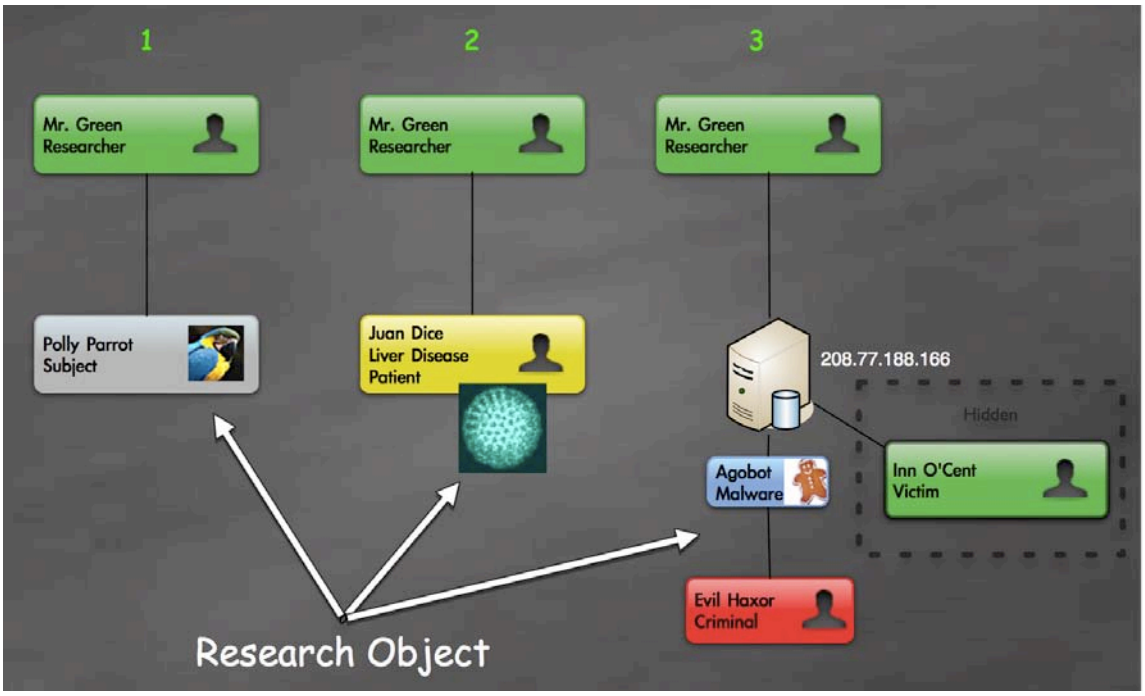

Figure 2.

In the illustration above, the term "research object" may be a more applicable

term than “research subject” if a researcher is trying to separate the focus of a research activity and the individual may be harmed.

Passive Research:

A researcher directly observes his subject in Context #1. These observations are passive and do not require direct intervention. If the observed behavior is non-controversial, in a public space, and no personally identifiable information is obtained, the risk to subjects is minimal or non-existent. Participants may not even need to provide consent.

Active Intervention:

In Context #2, a researcher studies disease organisms in the liver of a living human (i.e., the "research subject”). The organism is the object of the study. In order to observe the liver, the researcher must perform experimental procedure(s) on the human owner of the liver. This entails both informed consent (permission) and possibly some risk of harm to the human "subject." The researcher must directly interact with the subject. Obtaining informed consent is facilitated by a face-to-face interaction with the researcher, engaging in dialogue, asking questions, and actively participating in research procedures. It is clear who is at risk, what those risks are, and whether those risks materialize as harm.

The final situation, Context #3, indicates the studied object is either a piece of malicious software (e.g., a "bot"), or a criminal who infected a third party’s computer and is engaged in illegal activity. A researcher may directly intervene with the malicious bot software, but the owner and/or other users of the infected computer are completely

unaware of the bot software, the infection, and either the criminal or researcher's manipulation of their infected computer. The best identifier that the researcher has is an IP address. Some regulations consider an IP address to be an identifier. However, an IP address does not allow the researcher to identify or communicate with the owner and/or users of the infected computer any more than having a street address guarantees that the researcher can find and communicate with a specific individual at a residence.[xxiii] This makes obtaining informed consent impractical at best, or impossible at worst when the research may discover millions of infected computers' IP addresses. The risk that must be balanced is not risk to the object of the research (i.e., the bot), but to the humans who own and/or use the infected computer. If the researcher's manipulations cause the accidental destruction of data within the infected computer, the harm is indeterminate and possibly unquantifiable (yet to the person whose data is affected, significant and irreparable harm may occur).

If research involves passive observation, the risk to the humans may only consist of harms that occur from disclosure of their personal identity, behavior, or participation as a research subject. The risk of harm is revealing confidential information about the human. When the research activity is directly active, the harm can be physical or financial, even if a private is never exposed. Risks can include compromising the availability and integrity of information, along with the systems that contain information. The potential harms can expand beyond the primary owner and/or users of the computer system. Secondary harms are likely if individuals rely on the same information bases and/or information systems.

Situations 2 and 3 differ dramatically in terms of informed consent. In situation 2, the human has given consent, accepted the risk of participating in research, and is aware of possible causes of harm. In situation 3, harmed individuals cannot know the cause. They are unaware their computer is involved in research activity. Harm simply occurs and the harmed humans must guess the cause. It is easier for an individual to blame a criminal or buggy software when they suffer a computer malfunction. If a researcher causes harm the researchers and their institution may suffer public ridicule or reputational harm.

Spatial distance

The digital world can affect anyone connected to it no matter where they are in the physical world. The digital arena allows us to do wonderful things, but the harm inflicted through digital technology on companies that do not protect themselves from attack or unsuspecting individuals can be tremendous. We are all familiar with the concept of "distance" between two objects in physical space. Distance also manifests in a logical sense and a temporal sense.

Logical distance

The concept of logical distance has to do with the hierarchical relationships within ICT systems. These relationships form a graph with multiple intermediate levels. End users can be viewed as nodes at the edges of this graph. They obtain ICT services from platform or application providers, who rely on foundational computing resource service providers and network transport providers. All of these providers are interdependent and the failure of any one can cause failures in others, with the end result that an application

end users depend on does not work. The more layers there are in this provider/consumer hierarchy, the greater the logical distance between the researcher and the end users who may be affected by disruption of a service being subjected to an experiment by a researcher. In other words, an action taken by a researcher that affects a provider may affect a company that consumes the provider's services and sells other services to secondary providers, who have their own customers who may be affected by disruption to the upstream provider in a cascading manner.

Temporal distance

Software producers can use vulnerability knowledge to fix the problem and enhance reliability in the software system. The same knowledge can cause harm. Developing and testing patches, notifying customers or computer users, and downloading and implementing the patch takes time. In the gap, malicious actors can take advantage of those vulnerabilities to penetrate an unprotected system and either steal personal data or compromise the computer to distribute viruses.

Vulnerabilities affecting over 150 vendors were discovered in the Domain Name System in 2008.[xxiv] Public disclosure of the vulnerability occurred six months after the initial discovery. During those six months, vendors were notified and attempts were made to maximize the number of vendors who could provide a mitigating patch for the problem upon public notification. When the vulnerability was revealed, providers and consumers raced to patch their systems before malicious actors could exploit the vulnerabilities harmfully. Calculating vulnerability risk requires a harm function that sums harm over time and benefit over time.[xxv] When a researcher publishes vulnerability information, the

resulting harm may continue long after publication occurs, not just immediately afterward.

IV. Where Do We Go From Here? Next Steps.

We need to solve two fundamental problems to improve the situation we have just described. First, we must facilitate the identification of humans who could be impacted in some way by using ICT that is the subject of research, or is used by CS researchers as a conduit for interaction or intervention with humans. Without this comprehensive understanding of human involvement, potential causes of harm will be underestimated or be left out of the equation altogether. Second, we need to shift the focus of ethical review onto those situations in which there is a greater than minimal risk to humans from research activities themselves, not on whether there is direct intervention with humans as the subjects of research.

CS researchers want to do good for society, but good intenion alone may not be enough. Ethics can be the basis for conscious decisionmaking about research methodology that reflect one's intentions and their source of "consciousness, mindfulness, honesty, and sensitivity."[xxvi] Ethicist Annett Markham suggests self-reflective questions for researchers. "What is the purpose of this research project?" "What is the potential or desired outcome and why is research necessary to this outcome?"[xxvii] We can also ask "how would a stakeholder view my actions and interpret my intent? Would they feel grateful, neutral or resentful?"

A. Balancing risks and benefits of research using stakeholder analysis

Stakeholders include researchers, human subjects, society and criminals/attackers.

These broad categories can become: researchers and their programmers; vendors who use the internet to sell products or internet service providers (ISPs); the programmers for vendors; clients and customers of websites, online stores and ISPs; and criminals who exploit internet-based services and/or the data that they are able to discover through technological vulnerabilities. Some academic research activities can look like criminal activity, even though they exist to track an individual's reactions to attacks and not to exploit personal data (e.g., stealing credit card information).

Rarely do researchers consider *all* of the stakeholders affected by exploitation of a serious vulnerability when determining when/how to disclose vulnerability information and proof-of-concept (POC) exploit software/hardware, or justifying their actions. There is no widely recognized and adopted standard methodology in the CompSec community, however there are examples of this technique being used.[xxviii]

In several research cases involving electronic voting machines researchers have chosen to publicly disclose vulnerability information before notifying anyone. They feared a massive legal backlash from voting machine vendors. They believed it was necessary to inform the public so they could continue to trust the integrity of the election process. Tremendous political and financial stakes are at risk with the potential of voting fraud. Vendors are motivated to refute vulnerabilities, and researchers can gain positive publicity by exposing technological flaws. The public is the primary beneficiary of disclosure and mitigation of voting systems' vulnerabilities. Many other stakeholders are involved in the certification, purchase, operation, and monitoring of voting machines and

election results. There is little or no discussion of the positive or negative effects on other impacted stakeholders, such as:

- **Local municipal governments:** These government entities store, transport, and operate voting systems. They function on fixed annual budgets. They may be elected (accountable directly to the voters) or appointed (accountable to elected officials). They rely on volunteers who help operate polling locations.
- **State and federal legislatures**: Legislatures appropriate funds to purchase or maintain voting systems. Replacing voting machines is costly and it is time-consuming to appropriate funds that make these systems available or replace them. Help America Vote Act (HAVA) money was spent years ago and funding to replace voting systems it purchased for the states must come from future federal appropriations (unlikely in the current political climate), or state funds (which are similarly politically difficult to obtain presently).
- **State government executives** (i.e., Secretaries of State): State executives establish certification standards for voting systems. All systems must pass an evaluation based on existing (or newly defined) standards. Evaluations take months or years to complete.
- **Citizens** want clean and fair elections and trust in the voting process. Citizen watchdog groups monitor government activities and hold their representatives accountable for action or inaction. Private citizens volunteer as election monitors in many precincts to ensure that votes are properly handled and tabulated. These groups may assist in reforming the voting system, or may actively oppose elected officials who they believe do not ensure a trusted voting system.

- **Law enforcement agencies** at the state and federal level enforce voting laws. If a voting system is questioned so near an election reverting to paper ballots is impossible, tensions about election fraud may arise. Doubt about election validity can incur significant costs in both civil and criminal legal actions. In *Bush v. Gore* in 2000, civil legal process moved forward at "deliberate pace," which did not fit mandated time frames for recounts or challenges to election results.

It is challenging to identify the optimal means of notification that balances the benefits and risks to all stakeholders. There may be risks beyond the control of researchers and it may be difficult to identify an entity who will respond to reports of serious flaws in the voting system and who can assist in remediation or implementation of additional audit mechanisms during an election with flawed equipment.

The timing of vulnerability announcements plays a large role in potential harm that could occur if an election occurred using flawed equipment. Factors are the speed with which individual stakeholders can act, the fixed time frames of elections in the United States, and the balance of vulnerability information disclosed against the mitigation information that could secure the integrity of the voting system. Over time, trusted venues may exist in which researchers can publish sensitive information. Researchers may become more familiar with coordinated vulnerability disclosure mechanisms that are well positioned to coordinate sensitive disclosure and mitigation efforts.

B. Shift to regulating research that causes harm to humans rather than focusing on the harm specifically to human subjects of research

We propose a paradigm shift from assessing "human subjects research" to assessing "human harming research" in the CompSec research arena. The idea that direct interactions between individuals are the only source of research harm is obsolete in research fields that connect individuals and machines globally. Similarly, focusing only on data and identifiability of individuals who are the subjects of research leaves out other types of non-trivial harms.

The distance between technical expertise available to IRBs and how researchers portray risks in their research protocols is beyond the scope of this paper, but has been noted by others.[xxix] Borenstein explains that "[t]he goal of protecting human subjects should not be brushed aside merely so that researchers can proceed with their work. It should also not be forgotten that an effective review can detect flaws and prevent limited resources from being wasted."[xxx] The goal of an effective review of the technological issues cannot be achieved without available expertise on IRBs.

Recommendations to improve the efficiency of IRB review suggest allowing researchers to use established standards to bypass IRB review altogether.[xxxi] If researchers remove all data in the list of 18 HIPAA identifiers, current research can be exempt from review. Anonymizing data increases privacy and encourages expedited review of research. However the techniques used to anonymize data are generally not sufficient to prevent re-identification of data.[xxxii] The standards-based checklist only addresses privacy risks associated with data and ignores risks to integrity or availability of information or information systems.

ICT is constantly evolving, converging, and becoming more ubiquitous and transparent. As this occurs, ICT starts to be taken for granted the same way that those who live in the developed world rarely give a second thought to roads, electricity, or drinking water. The result is greater and greater chances that disruption or damage of ICT components will have human-harming effects.

We would like to begin an ideological shift toward regulating research that could cause harm to humans as an alternative to the current regulatory bent that focuses solely on the harm to *human subjects of research*. In the CS world, it is difficult to determine who is *actually* affected by CS research and it is better to consider the potential for harm when designing research and create a strategy to minimize the potential for harm and mitigate harm when it occurs. This also achieves the goal of optimizing ethical review of all research that involves ICT, including data-intensive studies in the biomedical research field, as the same type of analysis of stakeholders and potential benefits and harms to them are identical. We believe that over time, the capacity of IRBs to efficiently perform their valuable oversight duties, and ability of CS researchers to efficiently structure their research protocols and present them to IRBs for review, will bridge the distance and work toward consistent ethical analysis in CS Research.

Acknowledgements

The authors would like to thank Michael Bailey and Erin Kenneally for generously allowing their conversations and correspondence to help in writing this work, to participants in the Menlo Report working group for spirited debates about ethics in ICT

research, to Elizabeth Buchanan and Laura Odwazny for exploring the concept of "distance" as it applies to internet research under the Common Rule, and to Halle Showalter Salas for encouraging simplicity.

References

[i] Dittrich, D. and Kenneally, E., eds. Applying Ethical Principles to Information and Communication Technology Research. Forthcoming (2011).
[ii] 45 CFR 46. 102(d)
[iii] 45 CFR 46, 102(f).
[iv] Illinois Whitepaper, Gunsalus et. al., 2006 p. 9; http://www.primr.org/uploadedFiles/PRIMR_Site_Home/Resource_Center/Articles/11.%20Illinois%20Whitepaper.pdf; accessed October 2, 2011
[v] Buchanan E, Aycock J, Dexter S, Dittrich D. Hvizdak E. Computer Science Security Research and Human Subjects: Emerging Considerations for Research Ethics Boards Journal of Empirical Research on Human Research Ethics: An International Journal, Vol. 6, No. 2 (June 2011), pp. 71-83, at 72.
[vi] The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Ethical Principles and Guidelines for the protection of human subjects of research, April 18, 1979.
[vii] Ibid.
[viii] Dittrich, D. & Kenneally, E., eds. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research. Forthcoming (2011).
[ix] Merriam Webster Dictionary, 2011 online.
[x] Ibid. note iv.
[xi] Ibid. note iii.
[xii] Ibid.
[xiii] Ethics in Engineering, Martin, M.W., Schinzinger, McGraw-Hill, pp. 88-89, 4th Ed. 2005.
[xiv] Ibid. note i.
[xv] Dan Kaminsky. Black Ops 2008 -- It's The End Of The Cache As We Know It. In Black Hat Briefings USA 08, Las Vegas, Nevada, USA, July 2008.
xvi Katie Moussouris and Adrian Stone. Threat Complexity Requires New Levels of Collaboration. http://blogs.technet.com/ecostrat/archive/2009/07/27/threat-complexity-requires-new-levels-of-collaboration.aspx, July 2009.
[xvii] Microsoft Security Response Center. Coordinated Vulnerability Disclosure at Microsoft. http://go.microsoft.com/?linkid=9770197, April 2011.
[xviii] Jason Livingood, Nirmal Mody, and Mike O'Reirdan. Draft Recommendations for the Remediation of Bots in ISP Networks. http://tools.ietf.org/html/draft-oreirdan-mody-bot-remediation-03, March 2010.
[xix] Dittrich, D. and Kenneally, E., eds. Applying Ethical Principles to Information and Communication Technology Research. Forthcoming (2011).
[xx] Richard Domingues Boscovich. Microsoft Neutralizes Kelihos Botnet, Names

Defendant in Case.
http://blogs.technet.com/b/microsoft_blog/archive/2011/09/27/microsoft-neutralizes-kelihos-botnet-names-defendant-in-case.aspx, September 2011; accessed October 1, 2011.

[xxi] Kaspersky blog poll, http://www.securelist.com/en/polls?viewpoll=207796946

[xxii] David Dittrich, Michael Bailey, and Sven Dietrich. Towards Community Standards for Ethical Behavior in Computer Security Research. *Stevens CS Technical Report* 2009-1, April 20, 2009.

[xxiii] Ohm, Paul, Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization (August 13, 2009). UCLA Law Review, Vol. 57, p. 1701, 2010; U of Colorado Law Legal Studies Research Paper No. 9-12. Available at SSRN: http://ssrn.com/abstract=1450006

[xxiv] Sid Faber, Responsible Disclosure: A Case Study of CERT VU#800133, "DNS Cache Poisoning Issue," Domain Name System Operations Analysis and Research Center (DNS-OARC) Workshop, 2008.
https://www.dns-oarc.net/files/workshop-2008/faber.pdf.

[xxv] Ashish Arora and Rahul Telang. Economics of software vulnerability disclosure. *IEEE Security & Privacy*, 3(1): 20–25, 2005.

[xxvi] Ibid. note xx, p. 2.

[xxvii] Markham, AN. citation needed. Ethic as Method, Method as Ethic: A Case for Reflexivity in Qualitative ICT Research, *Journal of Information Ethics*, 15(2): 37-54, at 51 Fall 2006.

[xxviii] David Dittrich, Felix Leder, & Tillmann Werner, "A Case Study in Ethical Decision Making Regarding Remote Mitigation of Botnets," in Workshop on Ethics in Computer Security (WECSR '10), Tenerife, Canary Islands, Spain, January, 2010; Dittrich, D. and Kenneally, E., eds. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research.

[xxix] Illinois Whitepaper, Ibid viii; Borenstein, J., The Expanding Purview: Institutional Review Boards and the Review of Human Subjects Research, *Journal of Clinical Research Best Practices*, Vol. 5, No. 2, p. 5, February 2008; http://firstclinical.com/journal/2009/0902_AIR_Purview.pdf

[xxx] Ibid.

[xxxi] Department of Health and Human Services. Information Related to Advanced Notice of Proposed Rulemaking (ANPRM) for Revisions to the Common Rule, September 2011. http://www.hhs.gov/ohrp/humansubjects/anprm2011page.html

[xxxii] Ibid. note xxi.