Comments To HHS-OPHS-2011-0005: Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators

This response to the solicitation for comments to the ANPRM reflects the guidance proposed in the Menlo Report: Ethical Principles Guiding Information and Communication Technology Research (ICTR). The Menlo Report is an intellectual framework to address the need for ethical guidance in identifying unique issues, and understanding and applying principles in ICTR. It is rooted in the foundational principles espoused by the Belmont Report and codified in the federal human subjects regulations (the Common Rule). It is informed by a multi-year working group of community stakeholders in this dialogue and supported by the U.S. Department of Homeland Security.

The Menlo Report specifically addresses the computer science sub-discipline of information security research that involves studies related to ICT vulnerabilities, digital crime, and information assurance for critical infrastructure systems. These are areas where harms are not well understood yet are potentially significant in scope and impact. In response to this request for comments, we contend that (i) certain of the unique concerns raised by ICTR are not sufficiently reflected in these proposed changes, and (ii) that several of the proposed changes will impact ICTR in ways that do not advance the overall goals of improving the efficiency of the IRB process and enhancing protections for human subjects.

Submitted by: Erin Kenneally, Cooperative Association for Internet Data Analysis, University of California San Diego; Dave Dittrich, University of Washington; and, Michael Bailey, University of Michigan. These comments are the product of our individual work in the domain of information and communication technology research and authoring roles with the Menlo Report, and do not necessarily represent the views of our respective employers, the Department of Homeland Security, or the contributing members of the Menlo Report working group.

Sec. II. Ensuring Risk-Based Protections: Refinement of the existing risk-based regulatory framework

In general, our comments about refining the risk-based protections as it relates to ICTR are that the definitions of "minimal risk" (Question 1) and "reasonably foreseeable risks or discomforts" (Question 4) should account for the evolving context within which research is occurring. Specifically, calibrating the review process to the risk of research means broadening consideration beyond direct

human harms and protection of associated data, to include probable harms involving humans who are indirectly at risk of harm from interactions with ICT, as well as computers and networks that are the functional extensions of persons who may be impacted by research.

Revisions to evaluation and implementation of risk-based protections should bear in mind the changes that stem in large part from the attributes of ICT: scale, speed, tight coupling, decentralization and wide distribution, and opacity. Challenges identifying harms and benefits in ICT environments include: the scale and rapidity at which risk can manifest; the difficulty of attributing research risks to specific individuals and/or organizations; and our limited understanding of the causal dynamics between the physical and virtual worlds. This environment complicates achieving ethically defensible research for several reasons. It results in interactions with humans that are often indirect, stemming from an increase in either logical or physical ``distance'' between researcher and humans to be protected over research involving direct intervention. The relative ease in engaging multitudes of distributed human subjects (or data about them) through intermediating systems speeds the potential for harms to arise, and extends the range of stakeholders who may be impacted. Also, legal restrictions and requirements have expanded considerably since the 1980s, and ICTR is unquestionably subject to a variety of laws and regulations that address data collection and use.

While it is true that these individual complications are shared by traditional biomedical and behavioral research, the simultaneous confluence of these complicating factors occur with regularity in ICTR. Compared to our institutionalized or socially internalized understanding of harm related to physical interactions with human subjects, as a society we are relatively inexperienced regarding qualitative and quantitative assessment of damages and harms in the digital realm.

With regard to proposed mechanism for protecting subjects from informational risks, the suggestion to "[use] procedures which are consistent with sound research design and which do not unnecessarily expose subjects to risk". On its face, this is a laudable goal, however the term "subjects" needs to be broadened to include information and communication technology itself as the subject of research, not just direct interaction or intervention with humans. There are also implicit assumptions that there exist "procedures which are consistent with sound research design", which is not always the case with emerging fields of study of rapidly advancing information and communication technologies, and that technology researchers uniformly agree on what research procedures would "unnecessarily

[present] risk" and that IRBs have technical expertise available to them to provide either expedited or other simplified review mechanisms, neither of which are solid assumptions at this time. This is underscored by considering availability and integrity risks instead of the term "non-information risk" in Question 5.

Question 7: What research activities, if any, should be added to the published list of activities that can be used in a study that qualifies for expedited review?

- Research involving and/or focusing on ICT is often mistakenly dismissed as not involving risks to human subjects. We strongly advocate that there are several categories of ICTR that should not qualify for expedited review or be automatically excused from review, even if it conforms with the HIPAA Data Security requirements.

Question 11: What are the advantages of requiring that expedited review be conducted by an IRB member? Would it be appropriate to instead allow such review to be done by an appropriately trained individual, such as the manager of the IRB office, who need not be a member of the IRB? If not, what are the disadvantages of relying on a non-IRB member to conduct expedited review? If so, what would qualify as being ``appropriately trained''? Would the effort to make sure that such persons are appropriately trained outweigh the benefits from making this change?

- The advantage of allowing expedited review of ICTR by a non-IRB member is that IRBs are currently not familiar with the technical 'anatomy and physiology' of the ICTR platforms so their ability to issue-spot and evaluate the ethical soundness of this type of research is limited. Institutionalized guidance on the protection of research subjects has not kept pace with the rapid transformations in information technology and infrastructure that have catalyzed changes in research substance and mechanics. This gap affects both ICT-specific research (ICTR) such as information systems security, as well as biomedical and behavioral research enabled by the 'use' of ICT, insofar as both are confronted with research risks not envisioned in purely physical contexts. For these reasons, and owing to the attributes of ICT discussed above (scale, speed, tight coupling, decentralization and wide distribution, and opacity), we contend that outside review by a well-qualified domain professional may in fact be necessary. The qualifications of an appropriate reviewer is something that is entirely achievable and practicable, albeit a decision that will necessitate discourse and agreement

among ICTR community members.

Question 23: Should a request to waive informed consent trigger a requirement for IRB review?

- Yes. As detailed in our subsequent comment to Section IV. Question 43 there are justifiable reasons why it may be impracticable to obtain informed consent, and when information and communications technology is involved, the frequency of this occurring may be greater than in traditional research contexts. Nevertheless, in the interest of ensuring that the study's degree of risk is afforded the attention it warrants, the reasons justifying a waiver of informed consent in ICTR do not necessarily justify a waiver of oversight. Informed consent is a mechanism to heed the underlying Respect for Persons principle. To the extent that ICTR involves novel constructs that serve as the functional equivalent of informed consent or beg for a more context-appropriate application of the underlying principle (e.g., obtaining consent from service providers or other legally authorized representatives, considering the computer systems that impact persons who are typically not research subjects themselves, challenges identifying harms and benefits in complex technology environments), review by an oversight authority is essential to ensure that research is ethically-defensible.

IV. Improving Informed Consent: Both forms and process

Question 43: Are there additional circumstances under which it should be permissible to waive the usual requirements for obtaining or documenting informed consent?

- ICTR often involves hundreds or thousands of systems owned/operated by persons who may meet the definition of "human subject" (if the investigator obtains data about individuals through an interaction or intervention or obtain identifiable private information about individuals) and therefore trigger the protections of the Common Rule. Wavier of consent is expected and necessary for many studies in this field since obtaining informed consent may be impracticable, it may not be technically feasible to identify subjects, or it may interfere with scientific integrity of the results. We believe that these circumstances are adequately addressed under the current waiver criteria [45 CFR 46.116(d)]. While existing waiver rules provide sufficient flexibility and protection for current ICTR, as this field becomes more prevalent, ICT advances, and increasing data heightens the identifiability risk, we contend that there will be a

need to re-evaluate whether the current consent and waiver models are appropriate and effective. For instance, in ICT network contexts, is the current interpretation of individual informed consent an appropriate mechanism for the principle of respect for persons? Is it necessary to reinterpret the idea of "diminished autonomy" and advocate for the use of proxies (e.g., legally authorized representatives such as network service providers) provide consent on behalf of human end users who are impacted by ICTR?

V. Strengthening Data Protections To Minimize Information Risks: Establishment of mandatory data security and information protection standards for all studies that involve identifiable or potentially identifiable data

Question 54: Will use of the HIPAA Privacy Rule's standards for identifiable and de-identified information, and limited data sets, facilitate the implementation of the data security and information protection provisions being considered? Are the HIPAA standards, which were designed for dealing with health information, appropriate for use in all types of research studies, including social and behavioral research? If the HIPAA standards are not appropriate for all studies, what standards would be more appropriate?

- We recommend reconsidering adoption of the HIPAA standards for purposes of the Common Rule regarding what constitutes individually identifiable information, a limited data set, and de-identified information. Specifically, we recommend against the exclusive use of the current HIPAA Privacy Rule definition of "de-identified information." We advise a re-evaluation of the set of identifiers that must be removed for a data set to be considered ``de-identified'' under BOTH human subjects regulations and the HIPAA Privacy Rule. Also, while the HIPAA definition of identifiable data is helpful, we recommend against the exclusive reliance on it to provide sufficient protection against identifiability risks. HIPAA's definition of identifiable information (data that identifies an individual (18 specific identifiers) or data that provides a reasonable basis to identify an individual, as determined by a qualified statistician) narrows the scope of information that would be deemed identifiable from what the Common Rule provides,``information about behavior that occurs in a context in which the 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).''

Further, we recommend categorizing all research involving the primary collection of identifiable data as well as storage and secondary analysis of existing identifiable data as research involving identifiable information.

Question 59: Would study subjects be sufficiently protected from informational risks if investigators are required to adhere to a strict set of data security and information protection standards modeled on the HIPAA Rules? Are such standards appropriate not just for studies involving health information, but for all types of studies, including social and behavioral research? Or might a better system employ different standards for different types of research? (We note that the HIPAA Rules would allow subjects to authorize researchers to disclose the subjects' identities, in circumstances where investigators wish to publicly recognize their subjects in published reports, and the subjects appreciate that recognition.)

- The protections required to secure data that poses informational risk are insufficient because HIPAA rules are inadequate (e.g., it does not address the limitations on anonymization techniques) and the security standards are not well defined. For instance, there is lack of clarity about what constitutes data that poses an informational risk. Similarly, reasonable interpretations of what 'data either directly identifies or provides a reasonable basis for identifying an individual' is so expansive as to render this definition meaningless.

  Specifically, the term "informational risk", especially when used as a discrete category along side physical and psychological risks, is flawed. By only including those risks that, "derive from inappropriate use or disclosure of information, which could be harmful to the study subjects or groups," and by concluding that, "legal, social and economic harms [] can usually be viewed as variations on those core categories," results in two fundamental limitations in achieving the stated goal of, "[calibrating the safety and monitoring procedures] to a study's degree of risk".

  First, information and information systems are put at risk through impacts not only to confidentiality, but also to availability and integrity. Harm can occur from having one's valuable data altered or destroyed, or by rendering a piece of technology inoperable. As more research and experimentation is focused on information and

communication technology itself, especially that technology that is "live" and in use, risks arise of harm to information or information systems (i.e., technology) itself.

Second, a focus on "human subjects" and researchers having direct interaction with those humans can result in research being inappropriately judged as "excused" or "exempt" from IRB review when in reality there may be a greater than minimal risk of harm that could result from research activities to humans who are themselves not even the subjects of research, but who merely rely on technology for their everyday personal or financial activities. Rather than taking a "human subjects" approach to inclusion of research that poses "greater than minimal risk," it is more appropriate to evaluate research activities themselves as having a greater than minimal likelihood of being "human harming."

Question 61: Are there additional data security and information protection standards that should be considered? Should such mandatory standards be modeled on those used by the Federal government (for instance, the National Institute of Standards and Technology recently issued a ``Guide to Protecting the Confidentiality of Personally Identifiable Information.'')?

- We should harmonize standards and best practices for data security and information protection. HIPAA is helpful but not comprehensive or sufficient to address all risks and harms.

Question 62: If investigators are subject to data security and information protection requirements modeled on the HIPAA Rules, is it then acceptable for HIPAA covered entities to disclose limited data sets to investigators for research purposes without obtaining data use agreements?

- No, because HIPAA covered entities (including "business associates") have the backing of that law to ensure accountability. Even though protections may be modeled after HIPAA, that does not assure a legally enforceable backstop if those protections are not implemented and adhered to. A data use agreement would be necessary to extend that accountability, and, it could encompass other data use restrictions that are beyond the scope of HIPAA Rules.

Question 63: Given the concerns raised by some that even with the removal of the 18 HIPAA identifiers, re-identification of de-identified datasets is possible, should there be an absolute prohibition against re-identifying de-identified data?

- Yes (unless the purpose of the research is to assess the efficacy of a de-identification standard or practice).

Question 64: For research involving de-identified data, is the proposed prohibition against a researcher re-identifying such data a sufficient protection, or should there in some instances be requirements preventing the researcher from disclosing the de-identified data to, for example, third parties who might not be subject to these rules?

- Yes. For reasons admitted in this notice, in light of emerging technologies and the increasing availability of data that poses a risk of re-identification, researchers should be prohibited from re-identifying data that has been de-identified and should ensure that such prohibitions are transitive to those with whom they further disclose/share the data.

Question 65: Should registration with the institution be required for analysis of de-identified datasets, as was proposed in Section II(B)(3) for Excused research, so as to permit auditing for unauthorized re-identification?

- Yes.

Question 66: What entity or entities at an institution conducting research should be given the oversight authority to conduct the audits, and to make sure that these standards with regard to data security are being complied with? Should an institution have flexibility to determine which entity or entities will have this oversight responsibility for their institution?

- Institutions should have flexibility in choosing oversight entities. Institutions should be able to select external parties/organizations to conduct audits.

VIII. Clarifying and Harmonizing Regulatory Requirements and Agency Guidance: Improvement in the harmonization of regulations and related agency guidance

Question 74: If all Common Rule agencies issued one set of guidance, would research be facilitated both domestically and internationally? Would a single set of guidance be able to adequately address human subjects protections in diverse populations and contexts, and across

the broad range of research contexts (including biomedical, national security, education and other types of social and behavioral research)?

- A driving motivation behind the Menlo Report is to extend the ethical principles and applications developed in traditional biomedical and behavioral research to the "new" research contexts involving information and communications technology. As such, there is implicit support for uniform treatment of these issues. If this uniformity cannot be achieved without a single governmental set of guidance, or if achieving consensus across the entire Federal government will significantly impede and/or ultimately prevent timely issuance of guidance, then there may not be reasonable justifications for allowing guidance to be issued by each agency.