Ethical Research: Human Subjects and Computer Issues

To examine some basic Human Subjects and Computer Issues. My name is Gerald Maguire. This material was developed in conjunction with my colleague Dr. Ellen McGee, retired professor from Long Island University.

[slide 2]

The goals of this lecture are to define some essential terms with regard to research on human subjects. To point out some of the codes that are relevant to this type of research, especially with respect to Sweden. To examine some basic principles for ethical research and to present some ethical issues related to computing and the use of computing to collect, process, and access data.

[slide 3]

First, we begin with: What is research? The US Department of Health and Human Services defines research with human subjects as: "Research means a systematic investigation including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge" And "Human subject means a living person about whom an investigator (whether professional or student) conducting research obtains: Data through intervention or interaction with the individual, or Identifiable private information."

[slide 4]

So what's intervention? Well, the law goes on to define intervention: "intervention includes both physical procedures by which data are gathered (for example, venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes." And "Interaction includes communication or interpersonal contact between the investigator of the subject."

[slide 5]

The law also goes on to describe what is private information: "Private information includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information which has been provided for specific purposes by an individual and which the individual can reasonably expect will not be made public (for example a medical record)." So the "Private information must be individually identifiable (i.e., the identity of the subject is or may be readily be ascertained by the investigator or associated with the information) in order for obtaining the information to constitute research involving human subjects." So it's important to understand that it's about this individually identifiable private information.

[slide 6]

Now, there are a number of other codes covering research ranging from the American Medical Association through the World Medical Association Declaration of Helsinki, giving Ethical Principles for medical research involving human subjects. But it doesn't simply include medical research, it can include any research that involves human subjects, and it may even include activities and observations about these people. Such as in a chat room or forum or in an MMOG. Where people don't expect that they're being observed for the purposes of research. Whee the expectation is that they're carrying out something in private or with only a limited set of people able to observe even though there may be a record of this action and this communication.

[slide 7]

A very important code is the Nuremberg Code, which resulted from the trial of physicians involved during the activities of World War II, where human subjects were treated in very, very despicable fashions. The result is the first step of the Nuremberg code says that: "The voluntary consent of the human subject is absolutely essential." So without that, they shouldn't be involved. That means the prisoners of war, prisoners, minors, etc. Who can't choose and can't say no, shouldn't be subjects of human research. The experiments should yield fruitful results for the good of society that aren't procurable by other needs or methods. And they're not random and unnecessary. So that means that it has to be vital to society to carry out this research. Otherwise, you shouldn't be doing research involving humans. We could be doing it with modeling or other sorts of things, not using humans.

[slide 8]

"The experiment should be so designed and based on the results of animal experimentation and the knowledge of the natural history of the disease or other problem under study that the anticipated results will justify the performance of the experiment." You need to lay the foundations before you can carry out this experiment with humans, and of course, "The experiment should be so conducted as to avoid all unnecessary physical and mental suffering and injury."

[slide 9]

Five: "No experiments should be conducted where there is an a priori reason to believe that death or disabling injury will occur;" With one of my favorite clauses here: "except, perhaps, in those experiments where the experimental physicians also serve as subjects." So physicians who are willing to experiment on themselves, then they might be able to carry out an experiment which could result in death or disability but only of themselves. So they better believe in doing it; otherwise, they are not going to carry this out. Number six: "The degree of risk to be taken should never exceed that determined by the humanitarian importance of the problem to solved by this experiment." So the result is it has to be highly worthwhile to overcome; whatever degree of risk that there would be, otherwise it should be carried out.

[slide 10]

[7] Proper preparations have to be made. [8] The experiment only be conducted by qualified persons. [9] During the course of the experiment, the human subjects should be at liberty to bring the experiments to an end. So the result is they should be able to say: "No. I do want to participate anymore", and it should end then. And [10] doing the course of the experiment the scientist in charge must be prepared to terminate the experiment at any stage if he believes in the exercise of good faith, superior skill and careful judgment that continuation of the experiments is likely to result in injury, disability, or death to the experimental subjects. So in the years since that time, there have been a number of large studies that have been discontinued because they were double-blind studies and one set of people were receiving a placebo, i.e., something which had no effect, and others were receiving a drug that was hopefully to have a positive effect. The experiment was ended because it was believed that it was unethical to not be giving these people a drug that would help them when the risk had been shown to be sufficiently low, and therefore the experiment was terminated early.

[slide 11]

Now in Sweden, the governing act regarding human research is "The act concerning the Ethical Review of Research Involving Humans." This act states, "... research that is conducted using a method the aim of which is to affect these objects physically or psychologically. The fundamental is the research is approved only if it can be conducted with respect for human dignity and if human rights and basic freedoms are considered at all times. Human welfare is to be given precedence over the needs of society and science." And that comes from the CODEX, Rules & Guidelines for Research. Additionally, the law set up a set of regional ethics committees to which you submit an application to be examined to decide whether the research that you're proposing does meet the requirements from an ethical point of view in order to be carried out. There is also a Central (national) committee that has a supervisory and appeals role.

[slide 12]

Now, SFS 2003:460 (SFS is the standard" numbering of Swedish laws) was revised in 2008, and this version clarified a number of things. One, it said that research is now understood to be scientific experimental or theoretical work to gain new knowledge, and developmental work on scientific grounds, but explicitly excludes work performed at the undergraduate level at universities. And this is a very, very major change in the law. And the definition of handling personal data is redefined, they say that "Research involving handling of certain personal data shall from now on be examined regardless of whether human subjects give their informed consent or not." So the research that clearly involves a risk of human subjects, whether it be interviews, surveys, etc. shall likewise be ethically examined. So simply interviewing someone, asking questions of them that might cause psychological harm can't be done without having an appropriate ethical review of it. If the finding is that it causes harms not justified by the value of the work, then the work should not be carried out.

[slide 13]

Another law that is important to know about this Högskolelagen, the higher education act. And it has a section in it that says "The Swedish government this Furthermore added an important goal for higher education and research that may influence choices of scientific perspectives and problem:" In that this research "shall support a sustainable development that creates a good healthy environment for this and future generations, economical and social welfare and justice." And so, in this course, you're going to have a module on sustainability, and these issues about sustainable development for this and future generations are part of this thinking about sustainability in multiple facets, and you can read more about this.

[slide 14]

In the Swedish Research Council's booklet: "Good research practices", it says, "research ethics isn't static". So that means the things which might not have been considered ethical problems could become ethical problems and vice versa. "The ethical considerations in research," it states, "are largely a matter of finding a reasonable balance between various interests that are all legitimate." And "The harm and risks involved may vary considerably depending upon the disciplinary domain. Thus different kinds of research call for different kinds of considerations." So there are many people who would immediately say Ah, yes, this medical research involving humans has to go to an ethical review board, but it also may be the case that things that we would consider more of social science research, also have to go to the review board. So it's important to be aware of, as this guide says the ethics are not static. You can't just say, "ah, for my previous study, I didn't need to do it, I don't need to do it now". You need to think about the particular research proposal.

[slide 15]

And returning back to the principles of ethical research. We said there were four principles: non-maleficence, autonomy, beneficence, and justice. And that leads to a set of rules. You need to have a proper research design. You have to think about the scientific validity of what we're going. We have to make sure we have the free and informed consent of the human subjects. And there has to be a favorable balance of benefits versus the risk. And we have to explicitly consider the confidentiality and privacy aspects of the research that we're proposing to do.

[slide 16]

Now, informed consent requires that we need to describe to potential participants the purpose of the study, the procedures (including the estimated duration), the risks and benefits. Explicitly disclosing are we providing anonymity, are we trying to provide confidentiality or not. The name and the title of the researchers involved. Compensation, if there is any. And the fact that they have the freedom to withdraw from the research at any time. Only then can they voluntarily consent to participate in the study. So if you leave any of these out, then obviously you haven't provided them with informed consent.

[slide 17]

So what about anonymity and confidentiality. Well. Anonymous means that no one can identify who provided the information, and it turns out that this is actually really difficult to assure. And so there's this interesting paper by Arvind Narayanan and Vitaly Shmatikov, "How To Break Anonymity of the Netflix Prize Dataset" Where Netflix purposely anonymized (they thought) the dataset which they made available to researchers to help them provide better ways of matching choices of films with people and these researchers showed that they can actually find out who the individuals were in real life and which particular film that they have watched. Despite the major effort that Netflix made trying to anonymize the data. So it turns out anonymizing data is very hard. With regard to confidentiality, confidentiality means you guarantee that the information will not be identified with the particular person. Thus although you may know who the participant is. You have to ensure that this person's data isn't identifiable in any of your results. And it also turns out that this too is difficult, because there may be reasons why others have to examine your data. So you cannot guarantee complete confidentiality.

[slide 18]

There are a number of laws and a number of policies with regard to collecting personal data. For example, the US Federal Trade Commission's Fair Information Practice Principles regarding notice [slash] awareness, choice/consent, access/participation, integrity/security, enforcement/redress when it involves children. The OECD has Guidelines on the Protecting of Privacy and Trans Border Flows of Personal Data, in particular, limiting data collection, examining data quality, specifying the purpose for the data being collected, limiting the utilization of the data, security safeguards, individual participation, and accountability. And you can read about all of these.

[slide 19]

So, I mention one or more of the vulnerable populations, particularly children, and this means anyone under the age of majority. And in some countries, it goes further. So, for example, the U. S. Children's Online Privacy Protection Act (COPPA) forbids sites from collecting personal information from children under the age of 13 without the consent of their parents. This means if you're thinking of doing a survey, you need to make sure that the person who you might have as a participant is able to make the decision themselves legally to participate. If they are under the age of consent, they may not be entitled to make that decision.

[slide 20]

We also have a big problem today between privacy and security. We got overhead imaging, satellite reconnaissance. We have got closed-circuit television in London on every corner, and many other places cameras are all over. We have people taking genomic data and mining it. We have people monitoring the use of computers, network access. We have biometrics build into many devices providing access doors. And buildings with sensors, microsensors, data mining, and medical data, etc.

John Woodward, in his testimony before Congress on Privacy vs. Security: Electronic Surveillance in the District of Columbia, said when all these different technologies are interlinked the ability to collect information and the amount of information collected increases significantly, thereby significantly increasing potential privacy invasions. So, these separate pieces of information that may be in isolation - yes, it's okay to have the camera there, and it's okay to collect this genomic data. But now, when we start to put it together, we cause a major privacy violation for individuals.

[slide 21]

There are also very large scale information systems, such as many biobanks for tissues, samples, etc. are being collected and are available to other scientists for analysis. We have biometric databases, such as the European wide biometric database called EURODAC. The European Union Visa Information System (VIS), the Schengen Information System (SIS II). We've got centralized healthcare records and laws about them. We have the Swedish National Patient Summary, which individuals can opt-out of and even block access to the records, but we also have epSOS an EU-wide cross- border eHealth service is being introduced. All of these things are collecting information and making it available. And it's important that you understand what the limitations are and what must be done to protect an individual's data.

[slide 22]

So as we said, there's a conflict between security and privacy. EURODAC and all of these other things are attempts to promote synergy. Right, they want to increase the effectiveness of healthcare, provide better border security, provide better healthcare, to enhance interoperability, so a patient who is sick in one place but normally resides somewhere else will be able to go into the emergency ward, and they will have access to the patient's previous medical records. But all of that is at some cost in terms of privacy. Now Europe has the European data privacy directive that requires strict containment of data within the specific systems, places restrictions on access, the purpose for which the data is collected and used - must be defined and access must be proportional to the task and not collecting data which is inappropriate for the task. And the individuals have a right to withdraw their permission for the use of this data. The purpose for collecting the data is strictly defined, and the data collected must be proportional to the task, and individuals have a right to withdraw their permission for the data being collected. Sweden has a number of laws that are relevant to these. The first of these is The Personal Data Act (abbreviated PuL), the Patient Data Act, and the Public Access to Information and Secrecy Act, which defines what things can be secret and what things are not secret. "

[slide 23]

Now, some of you may know that Iceland a number of" years ago made their genealogical, medical, and health healthcare records combined with DNA information available to a company. So they

provided data on 70,000 individuals. Today that company: deCode Genetics, Inc. has information of both genetic data and medical information on 500,000 individuals, and their goal is to "discover genetic risk factors for common diseases". You can go and read about their code of business conduct and ethics. So they are very aware of the fact that how having this data and tying it to specific diseases and potentially cause risks for the individual, For example, an individual having certain genetic factors might not be able to provide insurance - if the insurance company is "ah! Nope, there is too big a risk you're going to have this disease we are not interested in insuring you, or if we are, the if we're going to do it for a very high price."

[slide 24]

There are a number of interesting issues about computational genetics, because, of course, once the information been aggregated and is no longer associated with the individual and was cross-referenced with medical and genealogical databases, it turns out the information doesn't enjoy legal protection any longer, because it's not personally identifiable - but did the individuals who volunteered their information to be used in such a system actually have adequate disclosure? Did they know what they were agreeing to? Could they comprehend that, because at the time they might not have known about these other uses and, of course, there is also the question of who owns the rights and the access to that data? So, if they are able to reverse engineer the genetic code - who has the right to be able to reduce copies?

[slide 25]

There are also some interesting questions about cloud computing and research ethics. As recently (today), people gather their data, and they do the processing of this data in the cloud computing environment. That means they need to think about the encryption for the transfer of the data to and from the computing system. They need to think about the encryption of the data when it is being stored. They have to think about what are their rights to be able to access that data. And if the data center is outside their normal legal domain, what is the proper data stewardship that's going to be carried out to protect data. And who has the right to be able to access it? And increasingly today, it is common that people are crowdsourcing their data analysis. They are saying, "Hey, here's all of this data", but now when someone processes it and says, "Oh! I found this very interesting thing in your data" should they be one of the authors on your paper. So Elizabeth Buchanan and Michael Zimmer wrote a very nice paper about this, and I encourage you to go and find more about it.

[slide 26]

As I mentioned previously, Heidi McKee and James Porter, in their book: "The ethics of Internet research", have done extensive analysis of looking at: What are the ethics of collecting data from online sites, chat rooms, etc. Can you use logos or someone's avatar or screen captures in your research? Can you quote what someone says when they're in their second life character or from a mailing list or something they said in the chat room? Or shouldn't you? There is an Association of

Internet Researchers (AoIR), and they have an ethics working group. You can read more about this. And there is also the International Journal of Internet Research Ethics, which you may find helpful. There are a number of other references that you can read. I encourage you to think about what you're doing and act in an ethical manner, especially if you're going to involve someone other than yourself.