Department of Health and Human Services
Office of the National Coordinator for Health Information Technology
NHINRFI@hhs.gov

The Electronic Frontier Foundation (EFF) and the World Privacy Forum ("the Submitters") respectfully submit these comments on the Request for Information (RFI) concerning the proposed National Health Information Network (NHIN). (69 Fed.Reg. 65599, Nov. 15, 2004).

Executive Summary

The Submitters counsel great caution in planning for the NHIN. As organizations that focus on technology and civil liberties, we believe that the proposed NHIN poses enormous risks to individual privacy that will not be easy to control.

Proponents of the NHIN argue that a network for sharing electronic health records (EHRs) will save lives and money. They may be right – but these goals will only be achieved if the NHIN is, from the outset, openly and carefully designed and implemented with privacy and security in mind. Privacy and security must be thought of as NHIN goals, not as constraints or trade-offs. Patients will not trust the NHIN if it does not protect their privacy and is not transparent. Patient care will suffer if the NHIN is not secure against attackers who can manipulate or deny legitimate access to patient data.

We are especially concerned that NHIN proponents understate or underestimate the difficulties of creating a system that must be trusted to gather, store, process and disseminate personally identifiable medical information about millions of individuals.

Advances in information technology, including technology for privacy and security, are real and ongoing. But there should be no rush when the most private information in our society is at stake.

Discussion

I. Response to General RFI (1)

In his executive order of April 27, 2004, President George W. Bush spoke of a "nationwide interoperable health information technology infrastructure." < http://www.whitehouse.gov/news/releases/2004/04/20040427-4.html> The EFF and WPF believe that at this early stage in development, it is most appropriate to model a National Health Information Network (NHIN) in functional terms.

To define the NHIN in functional terms means that fundamentally:

- An NHIN is an information system whose primary purpose is to support the delivery of health care to patients.

- The major functions of an NHIN are the secure collection, storage, processing and transmission of information.

II. Response to (7)

Question 7 asks about the privacy and security considerations, including compliance with HIPAA, implicated by the NHIN.

The first step toward answering this question is to recognize that patient information in the modern health care system is the most sensitive and private information about individuals in our society. Medical or health information alone is highly sensitive; the confidentiality of the doctor-patient relationship is fundamental to our way of life. Indeed, as more and more of our personal behavior has become the subject of social and political controversy, from drug, tobacco and alcohol use to sexual and reproductive or contraceptive behavior, medical privacy is even more important.

A 2004 incident highlights these tensions. The Department of Justice subpoenaed abortion patients' records from at least six hospitals in Michigan, New York, and Philadelphia. The DOJ request was immediately controversial, and was perceived negatively by patients across the country. Some hospitals did indeed turn over the requested records. A Chicago hospital, however, refused to release its records, stating that the hospital would "lose the trust of its patients" and that persons with "sensitive medical conditions may be inclined to turn elsewhere for treatment." See Anna Johnson, "Court upholds ruling blocking government from obtaining abortion records from Chicago hospital*,"* Associated Press, (March 26, 2004). In a system as large and as dynamic as the NHIN is slated to be, these tensions will be magnified by the ease and scale of the associated data flows.

The modern health care system, of course, covers more than medical data in the hands of physicians. Public health systems, scientific researchers, and even the marketing departments of hospitals and pharmaceutical companies have access to patient information. Meanwhile, private insurance and government benefits systems that pay for medical care include significant amounts of patient and patient household financial data as well. Thus, the modern health care system links patient medical data with much other highly personal information.

Accordingly, ensuring privacy and security must be one of the central issues in designing the NHIN. If we merely sought to move health information efficiently, the Internet already exists as a platform for highly efficient information dissemination. But neither the Internet nor the entities that use it can today provide the public with reasonable assurances of privacy and security. Indeed, the privacy and security problems that plague the Internet today are an object lesson in the importance of building privacy and security into the initial vision of the NHIN; privacy and security cannot be an afterthought.

Attention to privacy and security is especially important given that most discussions of a NHIN or large-scale Electronic Health Record (EHR) system presume the need for unique identifiers. By definition, the NHIN would thus constitute a large-scale ID system that likely would become another national ID system alongside the SSN and state driver's licenses. "Such large-scale systems pose significant privacy and security challenges," and "could easily result in inappropriate linkages among nominally independent databases." See Computer Science and Telecommunications Board, National Research Council, *Who Goes There? Authentication Through the Lens of Privacy* at 3, (2003).

We will analyze the privacy and security considerations from a substantive and process perspective, respectively.

A. Fair Information Practices

Our starting-point is the concept of fair information practices. Fair information practices are not invariant, but most codes of fair information practices include these elements:

1) The Principle of *Openness*: the existence of record-keeping systems and databanks containing data about individuals should be publicly known, along with a description of main purpose and uses of the data.

2) The Principle of *Individual Participation*: each individual should have a right to see any data about himself or herself and to correct or remove any data that is not timely, accurate, relevant, or complete.

3) The Principle of *Collection Limitation*: there should be limits to the collection of personal data; data should be collected by fair, lawful means; and data should be collected, where appropriate, with a person's knowledge or consent.

4) The Principle of *Data Quality*: personal data should be relevant to the purposes for which it is to be used, and should be accurate, complete, and timely.

5) The Principle of *Finality*: there must be limits to the uses and disclosures of personal data; data should be used only for purposes specified at the time of collection; data should not be otherwise disclosed without the consent of the data subject or other legal authority.

6) The Principle of *Security*: personal data should be protected by reasonable security safeguards against such risks as loss, unauthorized access, destruction, use, modification, or disclosure.

7) The Principle of *Accountability*: record keepers should be accountable for complying with fair information practices.

Compliance with these basic principles is especially important for a system like the NHIN, which will contain and disseminate highly personal information about millions of people.

B. Consent

A familiar and central concept in medical privacy is informed consent. Consent -- simply put, a patient's ability to say yes or no to data collection, use or dissemination in a meaningful fashion – will be of particular importance to the validity, fairness, and trustworthiness of the NHIN. The concept of meaningful consent needs to be deeply ingrained in the NHIN, and must be built in to any form of EHRs.

There are numerous areas of medical practice where informed, meaningful, and unambiguous patient consent are critical. Certainly, one of these is medical research. The past five years have been rich with case studies on the impact of the digitization of medical data and the wonders of its use in research. The Mayo Clinic, for example, boasts of a 4.4 million item genetic database. According to consistent statements from officials at the Mayo Clinic and IBM, its technology partner, patients' genetic data is being actively used for research purposes. See, for example, Bob Brewin, "IBM, Mayo Clinic Take Next Data Mining Step; DB2-based system contains records on 4.4M patients," Computerworld, (09 August 2004).
Recently, the Cleveland clinic has begun installing the same IBM system for use in genomic research on patients. See "IBM Partners With Cleveland Clinic For Better Medical Data Access,"
CMP TechWeb, (September 30, 2004).

According to Mayo officials, patient consent was received for the use of their genetic data in the Mayo Clinic/ IBM database. When seeking consent from individuals for the collection, storage and use of their genetic data, many if not most researchers ask that consent be granted for future unspecified uses. See for e.g., Callan, Benedicte and Gillespie, Ian "Biobanks: From health protection to data protection," OECD Observer (December 1, 2003).

Consent was granted for the Mayo use, but what did this consent look like? What was the context of this consent? How was this consent obtained? Is the consent revocable? Were the participants informed of the privacy risks not just to themselves, but to all of their relatives who have a genomic tie to that data? Were the relatives with these genomic ties informed of the research use of the data?

These questions are particularly important in light of the 2004 Stanford study "Genomic Research and Human Subject Privacy," by researchers Zhen Lin, Art Owen, and Russ B. Altman that concluded that genetic data, even when broken down and anonymized by best practices, were still highly identifiable. See Science Magazine, Vol. 305 at 183 (July, 2004.) Thus, in a genetic database, there is no such thing as anonymizing data as in a traditional protocol or otherwise. Given this, the issue of patient consent for the research use of genomic data and other patient medical data must be a core issue that is considered as the NHIN is constructed.

Because consent is a very challenging item to acquire in any meaningful way, it is likely that in some cases, some data in the NHIN will not be appropriate for medical research use and should not be subject at any time to medical reporting or use in research. Some things should not be able to be signed away by patient consent. Because patient data will now be digitized should not mean that all of this data becomes available for indiscriminate use as a research gold mine.

The issue of consent must also be carefully considered in relationship to the digitized data contained in EHRs and the NHIN. Boilerplate HIPAA consent granted at the time a patient enters treatment cannot and should not be considered valid consent for unrestricted future research use of the NHIN data. We recommend that a revocable patient consent architecture should be created, with tiers so that some data are unavailable for use, such as genomic data.

Given the ease of distribution or EHRs and the attendant privacy implications, single-request, blanket patient consent will no longer be appropriate in such a system. In an electronically mediated health records system, revocable, reviewable consent architectures may be built into the system from the beginning. In such a system, patients would have the right to review the e-dissemination record of their EHR data flows and make decisions about granting or withdrawing of consent based on what they have seen. Patient representatives could be appointed at key junctures throughout the infrastructure to attend to patient questions and concerns. And finally, due to the serious risks of misuse of patient data flows, patients should have a private right of action when their data – or consent – is abused.

Statistical health reporting is closely related idea to the use of digitized patient data in medical research. The ONCHIT Strategic Plan, *Goal 4: Improve Population Health* states: "Population health improvement requires the collection of timely, accurate, and detailed clinical information to allow for the evaluation of health care delivery and the reporting of critical findings to public health officials, clinical trials and other research, and feedback to clinicians," <http://www.os.dhhs.gov/healthit/executivesummary.html>.

We understand that the value of mandatory health reporting of certain diseases such as Tuberculosis and Influenza via the Public Health Information Network (PHIN) and the National Electronic Disease Surveillance System (NEDSS) generally create overall improvements in health. However, we are concerned that this reporting as identified in the strategic plan will grossly expand given the digital capabilities of such a system. In the pilot projects and in the hospitals that have digitized data, there is a consistent pattern of greatly increased data use for reporting. Sometimes, this reporting is done at a regional level, sometimes the reporting is done for internal use. See, for example, "Analytics Move to the Clinic," CMP Tech Web, (March 1, 2004).

We are also concerned that statistical health reporting, for the first time, will be inextricably tied to personally identifiable information. Within an ambitious unified system such as the NHIN, it will be extremely challenging to ensure that data reporting cannot be reconstituted and rematched to individually identifiable information. This is true even of data subsets, and is particularly true of genomic information.

Our organizations have much experience with seeing exactly this sort of data reconstitution occur in other settings. It will take a great deal of effort and forethought in the planning stages to avoid this substantial and serious problem. If there is a data lapse in health reporting of information, and this information goes to a commercial entity or becomes attached to an individual, widespread mistrust of the NHIN would ensue and the ultimate goal of increased health would not be achieved. Individuals should not avoid health care due to fear of recriminations due to exposure of a sensitive health condition.

C. Electronic Health Records

The National Academy of Sciences (NAS) has identified two general privacy and security concerns associated with EHRs:

(1) Error or abuse at the level of individual organizations.

(2) Systemic flows of patient information throughout the extended health care system, often without patient knowledge or consent.

See National Academy of Sciences, *For the Record: Protecting Health Information* 54 (1997) ("For the Record"). Although the former threat model generally receives more attention, the latter is probably the more important.

1. Threat model for NAS (1)

A simple threat model for (1) would take account of at least the following variables: attacker's motive; attacker's resources; attacker's initial access to target data. Within this general framework, at least five basic threat levels can be distinguished: insiders who make "innocent" mistakes and cause accidental disclosures; insiders who abuse their record access privileges; insiders who knowingly access information for spite or profit; unauthorized physical intruders; and vengeful employees and outsiders. See For the Record, at pages 56-61.

What are the general threats to individuals' privacy and safety?

- **Unauthorized or excessive disclosure of patient information:** Medical records contain perhaps the most private and sensitive information in our society. (e.g., AIDS and other STD information, depression, etc.) The creators of the NHIN system should be mindful that even the world's best-secured databases are accessed inappropriately for inappropriate uses of the data, often by trusted insiders. In a system of the scale of the NHIN, this risk of inappropriate insider access will be quite high. Therefore, mitigation of these risks will need to be considered in the preliminary architecture of the system, and should not rely strictly upon a “tiered access” card access system for data protection.

- **Discrimination or adverse treatment based on patient information**: (including genetic data) Many companies use health records in making employment-related decisions.

- **Function creep:** The availability of massive quantities of digital records will be a compelling set of data for medical researchers and for various analyses, from social analysis to Census use. Will insurers or marketers of pharmaceutical companies have access to any subsets of the data?

- **Incorrect data, incomplete data, out of date information**: We are all familiar with the stories of patients who died needlessly due to prescription drug interactions or a mistake that was easily preventable. While a digital system of records aids in handwriting-related problems, it will not cure human fallibility. Our concern is that the use of EHRs and the NHIN is being viewed as a panacea that will magically improve patient health by reducing error. But a quick reality check can be had by analyzing the credit reporting system. It is fully digitized, yet individuals find profound errors in their credit reports. This system will be no different. Robust controls – management and process—need to be installed to mitigate electronic errors.

- **Lack of transparency**: The EHR and the NHIN need to be 100 percent transparent to patients. Individuals with data in the system must have the right to inspect, correct, amend, challenge, and in some situations, remove their information on an ongoing basis. Individuals must also have the ability to know every entity that has accessed their personally identifiable information or data subsets.

- **Data matching patients’ medical records**: Due to the sensitive nature of these records, they should not be matched with for example, IRS records , Census records, or other government records. There will be a great temptation to use the rich supply of NHIN data to track deadbeat dads, individuals with outstanding arrest warrants, and even missing children via genomic tracking. Some uses of genetic data for these purposes is already occurring. See for example “Mother Finds Son by DNA Sampling,” Korea Times, (June 12, 2004). As uses of this medical data expand into the social arena, we believe individuals will begin avoiding treatment. To the extent this begins to occur, the overall health of individuals residing in this country will decline. This area of privacy risk correlates to the Strategic Plan’s Goal No. 4, *Strategy 1, Unify public health surveillance architectures,* which states: “An interoperable public health surveillance system is needed that will allow exchange of information, consistent with current law, between provider organizations, organizations they contract with, and state and federal agencies." Strategy 1 needs to be balanced with privacy rights and the goal of encouraging all individuals who need treatment to initiate treatment with appropriate health care providers. This will require strict limits on data matching.

- **Law enforcement use of medical data**: FOIA litigation by the Electronic Privacy Information Center uncovered that the FBI obtained 257.5 million airline Passenger Name Records following 9/11, and that the Bureau permanently incorporated the travel details of tens of millions of innocent people into its law enforcement databases via a data warehouse with a terabyte of data. < http://www.epic.org/privacy/airtravel/default.html>. Will the NHIN data be used for these kinds of purposes by law enforcement? If so, that will have a deleterious impact on civil liberties and health care. Individuals should not be faced with the prospect of having to either consent to subpoenas or the potential datamining of their medical histories by law enforcement authorities, or not receive treatment.

- **Unfettered use of patient data for medical research** : ONCHIT's Strategic framework Goal 4, *Strategy 3 Accelerate research and dissemination of evidence* states that "Information tools are needed that can accelerate scientific discoveries and their translation into clinically useful products, applications, and knowledge." <http://www.os.dhhs.gov/healthit/executivesummary.html>. This is an admirable goal, but it is one that needs to be tempered with strong privacy, consent, and opt-out provisions for the patient. As already stated in these remarks, when patient records are digitized, there is a track record of this data being used for medical research in ways that are harmful and pose risks for patient privacy. Genomic records are not able to be appropriately anonymized prior to use in medical research, for example.

- **Datamining**: Generally speaking, datamining patient data should not be a primary goal in the adoption of EHRs or of the NHIN. Because of the privacy risks associated with datamining of patient medical data, it should not be done without an open, vigorous debate about the privacy risks and stringent controls on data use. Already, some physicians are finding that electronic medical records are a boon for improving "financial performance" through datamining patient records. "The advent of the electronic medical record has shifted the dynamic of data mining from merely a pharmaceutical marketing strategy to a powerful tool for clinical and financial performance improvement. Medical practices that have invested in EMRs are gleaning important knowledge from their own rich repositories of clinical information to identify populations for disease management, benchmark against quality standards, improve billing processes and track outcomes." See Neil Versel, "A beautiful mine : EMR adjusts data focus to include clinical, financial benefits," Modern Physician, 26 Vol: 7 Num: 7 (July 1, 2003). While cost savings is a legitimate goal, cost savings gleaned from extensive datamining should be balanced against patient privacy concerns and issues.

- **Unrestricted use of RFID**: RFID, or radio frequency identification tags, have not been specifically mentioned in conjunction with the NHIN yet. RFID is a threat to privacy in the healthcare context, especially when implanted in human beings via direct chip implant or via medical devices such as pacemakers. RFID use needs to be considered carefully in the NHIN context and should be curtailed to mundane uses, such as on prescription drug bottles. The use of RFID tags implanted on or within human beings in conjunction with an EHR needs to be avoided altogether.

D. Process issues

We have identified three primary issues related to process in regards to the NHIN:

- developing concrete threat models
- modeling information flows and points/streams of vulnerability
- privacy/security analysis of flows, vulnerabilities

1. Countermeasures are difficult

Privacy and security are both, in a very real sense, matters of economics. Thus, privacy and security countermeasures generally work by raising the costs of "bad" behavior. The two main complementary strategies are deterrence and prevention or imposition of obstacles. Tactics of deterrence include liability or accountability rules that typically look backward. Tactics of prevention include access controls, cryptography, and de-identification.

Unfortunately, today's records systems are poorly protected. Two recent examples illustrate this point -- more can be adduced.

For example, recently, a hacker penetrated T-Mobile systems, gaining long-term access to individuals' data. According to SecurityFocus, "An intruder had access to customer records, Social Security numbers,and private e-mail for at least a year ending last October. Trophies included celebrity cell phone snapshots, and a trove of restricted Secret Service documents stored on an agent's Sidekick." http://www.securityfocus.com/news/10271.

In another recent example – and one common to databases containing personally identifiable information – a secured George Mason University server with ID cards was accessed by hackers.

> From: "Joy Hughes, Vice President for Information Technology" <announce@gmu.edu>
> Date: Sun, 09 Jan 2005 14:08:49 -0500
> To: ANNOUNCE04-L@mail04.gmu.edu

> To: Mason Community
>
> From: Joy Hughes, Vice President for Information Technology
>
> Subject: Illegal Intrusion into University Database
>
> The university server containing the information relating to Mason's ID cards was illegally entered by computer hackers. The server contained the names, photos, social security numbers and G numbers of all members of the Mason community who have identification cards.
>
> The intruder installed tools on the ID server that allowed other campus servers to be probed. An Information Technology Unit staff member noticed the attack while reviewing system files as part of the university's internal controls procedures, and traced it back to the ID server. The compromised ID server was disconnected from the network and is no longer accessible. The police are currently investigating the break-in. The university is subject to dozens of probes and attacks each day.

The situation is even worse for patient information, which is uniquely sensitive and in many cases will have serious consequences if wrongly accessed or disclosed. Second, the time frame for protecting the privacy of medical data is at least the life of the patient, if not longer. A person can get a new SSN, bank account or unlisted phone number; no one can get a new body. Accordingly, when EHRs are at issue, the balance between deterrence and prevention must tip sharply in favor of prevention. This means that EHR information security cannot merely defend against today's state-of-the-art attacks; it must anticipate attacks that have not yet developed. Encryption that is adequate today is highly unlikely to be adequate in 20 years.

Furthermore, it must be recognized that deterrence and prevention can conflict in practice. Some discussions of security in the NHIN emphasize the need for accountability and thus the role of authentication. But "[t]he use of authentication when it is not needed to achieve an appropriate level of security could threaten privacy. . . . There is an inherent tension between authentication and privacy, because the act of authentication involves some disclosure and confirmation of personal information." Who Goes There at 5-7 (finding that "[a]uthentication can affect decisional privacy, information privacy, communications privacy, and bodily integrity privacy interests"); id. at 178 ("In the case of government, respecting the legitimate function of anonymity is even more crucial.").

We are also concerned that NHIN proponents underestimate the difficulties of designing and implementing such a large-scale network as well as its necessary technical countermeasures. A recent report of the FBI's difficulty with the Virtual Case File component of its Trilogy information system is a good example. This $170 million network, intended for the FBI's roughly 12,000 agents, is "over budget and behind schedule"; the Justice Department's inspector general has concluded that the system will not work and cannot be put into use. Among the problems cited was "difficulties in developing a system that is both secure and accessible to investigators." Eric Lichtblau, "FBI may scrap vital overhaul for computers," The New York Times at A1, A14 (Jan. 14, 2005).

Given the difficulties with designing a secure and accessible network for 12,000 users, a Markle Foundation report on privacy and security for EHRs appears far too optimistic in its discussion of access controls for the NHIN. According to the report, "[c]ontrolling access to electronic medical information using extensive specifications of an individual's role and functions is feasible. This type of system can offer a highly refined level of control over data access." Markle Foundation, *Connecting For Health: A Public-Private Collaborative* at 8 (Report and Findings of The Privacy and Security Working Group, June 5, 2003) ("Markle PSWG Report").

By contrast, a National Academy of Sciences report discussing the privacy and security implications of nationwide ID systems is considerably more pessimistic about multilevel access controls. "A further complication would result if it were decided that different users should be granted different levels of access to the database . . . . This raises query capability, access control, and security issues. . . . It is a significant challenge to develop an infrastructure that would allow multiple kinds of queries, differing constraints on queries (based on who was making them), restrictions on the data displayed to what was needed for the particular transaction or interaction, and varying thresholds for security based on the requirements of the user." Computer Science and Telecommunications Board, National Research Council, *IDs – Not That Easy: Questions About Nationwide Identity Systems* at 42 (2002).

It is useful to consider the Veteran's Affairs hospitals as a case in point. These hospitals have moved to a form of the new contactless Federal ID for managing health access, with plans to move to the full new standard quickly. See, for example, "Department of Veterans Affairs to use IWS EPI Builder for AAIP smart card initiative," Federal Computer Market Report at 4, Volume 28 Issue 16 (Sept. 6, 2004). The VA model is similar to the Markle proposal of tiered access, with that access being built into a single card standard. Unfortunately, the National Academy of Sciences warnings ring true in this implementation, which is highly flawed. First, the card itself is problematic. It contains numerous technological flaws, and presents a series of privacy risks related to its design features. Further, the backend management of the data is highly flawed. And the cost associated with the biometric capture for the card and conducting the background checks to assure tiered access is quite high. This model, if chosen, would not scale up to the degree necessary for use in the private sector.

Another of our concerns is that discussions of privacy and security in the NHIN seem focused on technical countermeasures with insufficient consideration of the human factors involved. For example, the Markle study of privacy and security quoted above only recognizes as an afterthought that its access control model may be “complex and cumbersome” for users. Markle PSWG Report at 35.

The Markle report also appeals to the use of hardware tokens, passwords and a public-key infrastructure (PKI) that relies heavily on cryptography. Again, none of these technical countermeasures will be effective without significant attention to the role and capabilities of the system’s users. See, e.g., Who Goes There at 5 (noting that while public-key cryptography “is often touted as a virtual panacea for . . . authentication and confidentiality challenges,” “implementation and deployment details are key to . . . effectiveness, security, usability, and privacy protection.”).

These “human factors” considerations are especially serious given the size and breadth of the population of likely NHIN users. A recent Kaiser Family Foundation survey found that only 31 percent of seniors (age 65 and older) have ever gone online, and that the poorer seniors were less likely to have been online. See Verne Kopytoff, Seniors slow to get health data online, S.F. Chronicle (January 13, 2005), <http://www.sfgate.com . . .>

2. Threat model for NAS (2)

While the challenges for threat model (1) are daunting, threat model (2) may pose even greater problems. As we have already noted, there are massive flows of information within the modern health care system to a wide range of users. Primary users include caregivers and their support system, like pharmacies and clinical laboratories. Secondary users include payors, insurers, government benefit agencies, accrediting organizations, bureaus of vital statistics and health departments, scientific researchers, as well as marketing firms and vendors of health-related products. See generally For the Record, at 73 (Figure 3.1) (flow chart of a hypothetical patient’s health information).

The risks of this massive information flow throughout the health care system stem from the widely divergent interests of the various users, particularly regarding the proper trade-off between the patient’s privacy interests and the users’ institutional or business interests. The fundamental problem is that health information “has business value to organizations such as insurers, employers, providers, and drug companies. Id. at 80.

Thus, “[w]ithout industry-wide standards or regulations governing the uses of health information by primary and secondary users, the information can [be] – and sometimes is – employed for purposes that violate patient privacy or are detrimental to the interests of the patient.” For the Record, at 77. The report gives an example of discrimination resulting from access to genetic information, such as loss of employment, loss of insurance coverage, or ineligibility for insurance. In some cases, discrimination was based merely on evidence of predisposition to a future occurrence of treatable diseases such as hemochromatosis and Huntington’s disease. Ibid. In another example, a pharmaceutical company acquired a drug reimbursement service, which used information in its database in order to try to convince doctors to prescribe the company’s drugs. The patients’ “privacy interests were compromised when confidential information about them was furnished to individuals (pharmacists) who were unconnected with the conduct or quality of their care. The lesson here is that "the interests of patients may not be well served by wide dissemination of health care information." Id. at 78.

3. ID Systems and Access Schemes for Patient Records

Finally, concerns about "systematic sharing" of EHRs are intertwined with efforts to establish an ID system for indexing patient records. The key issue here is the "linking" threats: as the National Academy of Sciences noted, "What generates the largest amount of controversy is the prospect that a universal health identifier will facilitate attempts to link information within and across much larger boundaries." For the Record, at 79.

In this context, it is important to again consider function creep; that is, (noting that "[t]he use of any single identifier as a universal identifier could expand beyond its initial intent and become widespread in other domains, just as use of the SSN expanded well beyond the realm of identifying Social Security records." Ibid.

The scale of the NHIN means that "if a single identifier is relied on across multiple institutions, its fraudulent or inappropriate use . . . could have far greater ramifications than if used in only a single institution." Who Goes There, at 10. Thus, the report recommended that "[a] guiding principle in the design or selection of authentication technologies should be to minimize the linking of user information across systems unless the express purpose of the system is to provide such linkage. Id. at 11. Moreover, such systems "should be designed to make secondary uses difficult, because such uses often undermine privacy, pose a security risk, create unplanned-for costs, and generate public opposition to the issuer." Ibid.

The ONCHIT Strategic Framework, *Strategy 3. Coordinate federal health information systems*, states*:* "There is a need for federal health information systems to be interoperable and to exchange data so that federal care delivery, reimbursement, and oversight are more efficient and cost-effective. Federal health information systems will be interoperable and consistent with the national health information network."
< http://www.os.dhhs.gov/healthit/executivesummary.html>

It is a feasible goal for the Federal health information systems to be interoperable with the NHIN. However, good interoperability does not mean that access for these two systems should be identical, specifically, the private sector should not be adopting the Federal ID card or its specifications as its means of access to the NHIN. Adoption of the Federal ID or its specifications for private sector NHIN access would be to encourage the adoption of a form of national identification by the private sector, and would be onerously costly to implement. Additionally, many extremely problematic issues relating to background checks would come up in a system based on the Federal ID card.

4. International Health Data flows and the NHIN

Several key issues revolve on the question of interoperability on a larger scale than the ONCHIT framework considers -- one that crosses U.S. borders and interacts with other nations. Issues of data protection differences between the countries, data interchange issues, and transborder data flows will impact any version of the NHIN that is implemented. These issues need to be taken into account and planned for from the beginning of the NHIN planning process.

Regarding data protection statutes, ONCHIT has declined at this point to consider the impact of the NHIN on citizens of other countries who may have greater statutory data protection than citizens of the U.S. This would include, for example, the European data privacy laws generally afforded under EU Directive 95/46 <http://europa.eu.int/comm/internal_market/privacy/law_en.htm >, and the Canadian privacy laws enacted under Canada's PIPED Act <http://www.privcom.gc.ca/legislation/02_06_01_01_e.asp >.

Foreign visitors to the U.S. who find themselves needing medical treatment should not be subject to data privacy issues that are at odds with the laws of their country, which in some cases may apply. For example, in the business community, there is a well-known anecdote regarding the fact that General Motors was not able to send a global phone book out due to the privacy provisions afforded by EU 95/46. See "For Your Eyes Only -- Europe's New High-Tech Role: Playing Privacy Cop to the World --- U.S. Companies Run Afoul Of EU Laws on Sharing And Collection of Data --- GM's Phone-Book Odyssey," David Scheer, The Wall Street Journal, (October 10, 2003). Medical data is much more sensitive than a phone number, and therefore this entire topic needs to be considered much more seriously and substantively than it has been to date. Data matching issues come into play here. How will collections and use of genomic data of foreign visitors be handled? How will consent be handled? And will there be data matching of foreign visitor medical data with other government data collected on foreign visitors, for example, travel data? What recourse will visitors to the U.S. have to opt-out of data collection during medical treatment procedures?

Additionally, transborder data flows and the privacy issues inherent in such transfers also need to be analyzed in regards to the NHIN and EHR systems. U.S. citizens traveling abroad often need medical care, and this care may be of a critical nature. Medical professionals in countries across the globe will need to have a practical, efficient way to access U.S.-based patients' medical records, digitized or not. This sort of interoperability must be considered in the NHIN and EHR system so that U.S. citizens and other participants of the NHIN do not find themselves receiving poor care in other countries due to a system that is inoperable, disliked, or distrusted in the rest of the world. We understand that ONCHIT does not have the responsibility to develop a global system. But ONCHIT does have the responsibility to develop a system that will not get in the way of what other countries are doing, and will not interfere with individuals' health care as they travel or work abroad. Given our highly mobile society, this is of critical importance.

5. Building Privacy Into the NHIN and EHR Architecture

To build privacy into the proposed digitized health network in a systemic manner, we recommend several key actions. First, a balanced approach needs to be taken that involves not just industry and government representatives, but also specifically and meaningfully involves privacy and civil liberties groups from around the country. For example, the ONCHIT strategic plan mentions a "Health Information Technology Leadership Panel." This panel needs to contain a balance of industry, technology, privacy, and civil liberties groups. Ideally, this panel should also include input from overseas privacy organizations, and should include input from privacy commissioners of other countries where applicable.

ONCHIT has not yet discussed what System of Records is being considered for the NHIN in regards to the implementation of the Privacy Act of 1974. It is very early yet; however, these kinds of questions about what kind of System of Records is being created are key tools in thinking through the privacy implications of an electronic record-keeping system. The time to consider Privacy Act implementation is at the beginning of a project, not the end, and the questions of privacy should be considered by all players, not just those specifically assigned to privacy.

Another tool along these lines is the Privacy Impact Assessment, which is mandated by provisions in the e-Government Act of 2002. Again, we recognize that it is very early in the process. However, the Privacy Impact Assessment is a critical thinking and analysis tool that is extremely helpful for government agencies to utilize at the beginning of a process of acquiring new technologies. The NHIN most certainly would benefit from an early, assertive, incremental PIA that is published publicly.

Additionally, we request that ONCHIT hold public hearings that focus on privacy issues, including the implementation of HIPAA, the Privacy Act, and the E-Government Act of 2002 (Privacy Impact Assessments) early in the process.

We are hopeful that ONCHIT will actively reach out to the privacy and civil liberties community for input, leadership, and guidance from the very beginning of this project all the way to its final implementation and eventual auditing and assessment.

III. Response to (8)

Question 8 asks how the NHIN framework could address "public policy objectives for broad participation, responsiveness, open and non-proprietary interoperable infrastructure."

Our review of the open literature on the NHIN leads us to be very concerned that the NHIN process will be dominated by large organizations that do not represent individuals' interests. There can be little doubt, for example, that enormous amounts of money are at stake here, and that major corporations in the health care and information technology industries have great financial interests in the design and implementation of the NHIN.

EFF and WPF are also concerned that the government itself has interests in the NHIN that may distort the public interest. For instance, Bio-Surveillance or Bio-ALIRT (Bio-

Event Advanced Leading Indicator Recognition Technology) was a technology project associated with the controversial "Total Information Awareness" program. We are therefore concerned as well that the NHIN development framework avoids being distorted by law enforcement or counter-terrorism interests.

Therefore, part of the answer is to use an open, public process that permits full participation by individuals, patients, consumers, and public-interest organizations that represent individuals, patients, and consumers. All relevant information should be made publicly available, and the process must avoid being obstructed by overbroad claims of the need to protect critical infrastructure information and proprietary information.

Another part of the answer is to take seriously the potential of open-source software, as opposed to proprietary software. For present purposes, a simple definition will suffice: open-source software is software distributed under a license that guarantees the right to read, redistribute, modify, and use the software freely. For a more complete definition of open source, see "Open Source Initiative OSI – The Open Source Definition," <http://www.opensource.org/docs/def_print.php>.

We believe that open-source software offers significant interoperability and privacy/security benefits for a NHIN, as your agency apparently recognizes in speaking of a "non-proprietary" interoperable infrastructure. It's critical to recognize that for large-scale networks, interoperability, privacy and security are not merely static goals or values; they are dynamic processes. For each, the relevant state of the world or situation will always be changing. Interoperable with what? During what time frame? As the network grows and new applications and users are added, interoperability needs as well as attackers' capabilities and incentives will evolve.

For this reason, ready modifiability of source code is essential. A key concept in open-source software is that code evolves, and that open-source processes make code evolution easier. The mere ability to read or view the source code is necessary but not sufficient; rapid evolution requires that users be able to experiment with and redistribute modifications. Indeed, Microsoft has recognized that open source "encourages a *process* around a core code base and encourages extensions to the code base by other developers" and that this methodology or process "promotes rapid creation and deployment of incremental features and bug fixes."

A. Interoperability

Basing the infrastructure of a NHIN on open-source software is more likely to promote interoperability than basing it on a proprietary platform. Because the source code and APIs for the "backbone" or "skeleton" of the network will be open and public, both open-source and proprietary software vendors will be able to write software on a level playing field for specific applications. The obvious analogy here is to the Internet, which has been designed around public, open protocols. The Internet has grown so rapidly and become a fertile ground for innovation in communications precisely because everyone knows how to design applications and devices to interoperate with the Internet.

By contrast, commercial vendors tend to only provide interoperability when and where they absolutely must do so; indeed, in the general software market, commercial vendors often deliberately break or obstruct interoperability in order to exploit power or dominance in the market. Given the broad scope of the NHIN, it should be expected that if its infrastructure were proprietary, the vendor of the proprietary software would make it hard for other vendors to interoperate. To use an antitrust concept, the NHIN infrastructure would be akin to an "essential facility" under one vendor's control. It cannot be in the public interest to have the NHIN infrastructure be controlled by proprietary software vendors.

B. Privacy and security

Open-source software also avoids the problem of "security through obscurity." Good privacy and security threat models assume that attackers are more motivated and persistent than defenders; attackers are far more likely to find the holes whether the source code is open or proprietary. Moreover, relying on the closed nature of the software creates two additional problems: a false sense of security is created, and it is harder to distribute trustworthy fixes when a hole is eventually revealed.

At the end of the day, of course, software security is a function of implementation quality. Reliability, security, bug density, feature-richness, appropriateness, cost, etc. must all be taken on a case-by-case or even component-by-component basis. Thus we cannot say that open-source software is necessarily more secure or more privacy-protective than closed-source or proprietary software. The only fundamental or necessary difference between open and proprietary software is its openness. But this single difference has significant implications for privacy and security.

If one really cares about privacy and security, one must extensively review and audit. Only open source software makes that possible – and in a publicly accountable way. NHIN participants and users will not need to rely on the representations of a proprietary software vendor; the actual source code will be available for inspection and testing.

Open-source software also amplifies available support resources. As the MITRE Corporation has noted, "the publicly available source code . . . enables many vendors to learn the platform and provide support." MITRE also observed that open-source software is debugged faster. "For example, the Linux community developed a fix for the TearDrop IP attack in less than 24 hours after it first surfaced on the Web." More generally, "[o]rganizations do not need to rely on a commercial provider's schedule for fixing a work-stopping bug but can, instead, opt to fix the problem themselves."

One can also argue that open-source software has other advantages due to its development process. For instance, because open-source developers have lesser or different business models, they may be less likely to distribute software with compromised privacy or security in order to reap short-term economic gains. Indeed, many open-source projects seek to enhance privacy.

In short, open-source software promotes software reliability and quality by supporting independent peer review and rapid evolution of source code. For a detailed report on the value of open-source software, see MITRE Corporation, A Business Case Study of Open Source Software (July 2001), available at <http://www.mitre.org/work/tech_papers/tech_papers_01/kenwood_software/kenwood_software.pdf>.