P1 P2 P3 P4 P5 P6 P7 P8

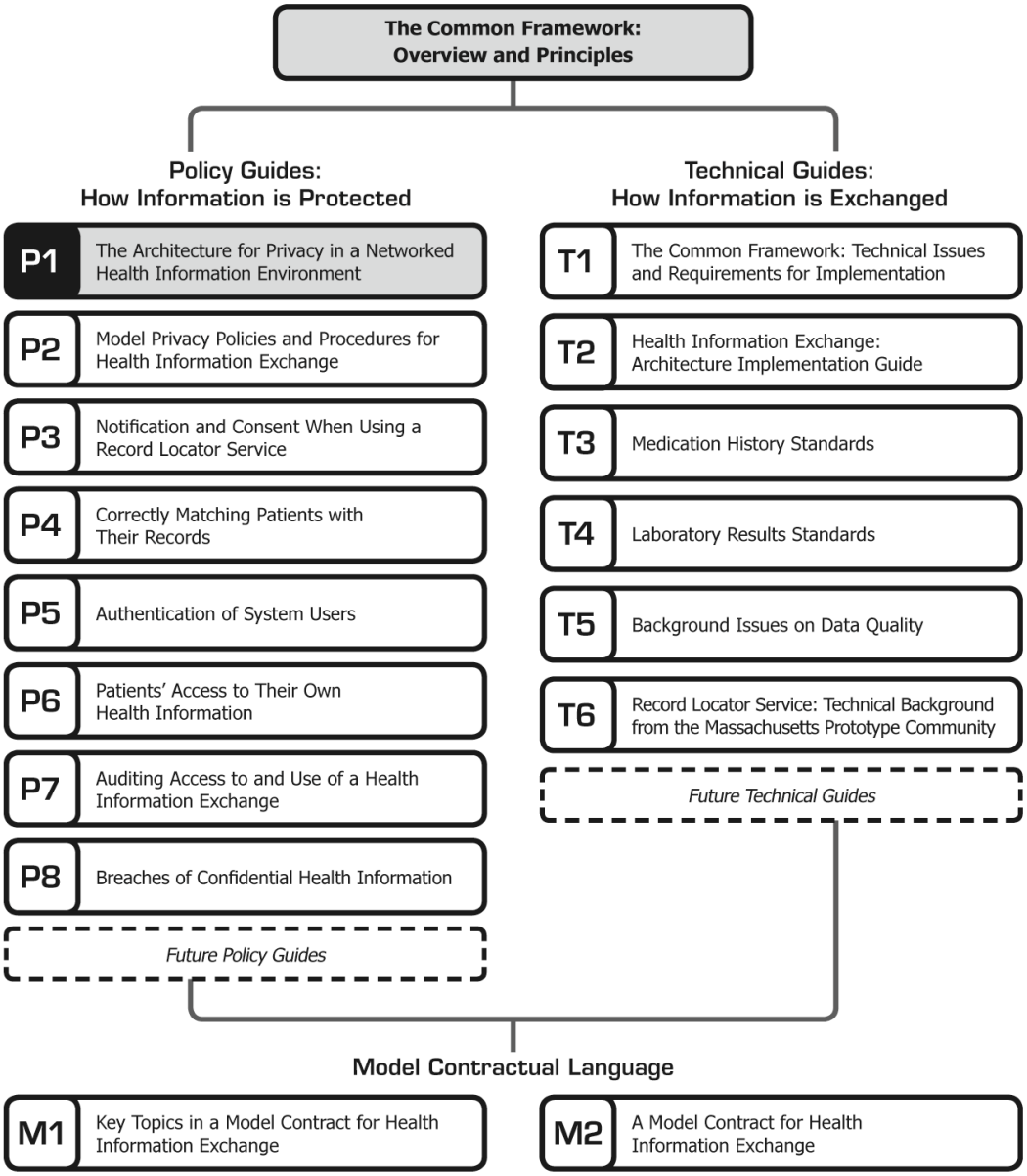
T1 T2 T3 T4 T5 T6 M1 M2

The Architecture for Privacy in a Networked Health Information Environment

The Architecture for Privacy in a Networked Health Information Environment

CONNECTING FOR HEALTH COMMON FRAMEWORK

The document you are reading is part of *The **Connecting for Health** Common Framework,* which is available in full and in its most current version at: http://www.connectingforhealth.org/. The Common Framework will be revised and expanded over time. As of April 2006, the Common Framework included the following published components:



The Architecture for Privacy in a Networked Health Information Environment*

Executive Summary

Introduction and Overview

A networked health information-sharing environment has the potential to enable decision support anywhere at any time, improving public and individual health, and reducing cost. Consumers and patients can benefit directly when their personal information is available to health care providers, and indirectly when their information is available in the aggregate to researchers seeking new ways to prevent, manage, or cure health problems. At the same time, the potential benefits must be weighed against the risks of privacy and security violations, which may increase if not addressed at the outset.

The accompanying document begins from the premise that any new health network needs to take into account the potential for such violations, and to build privacy and information security into its architecture from the outset, not as an afterthought. The document provides background on the issues at stake, explains the current status of health privacy, considers new challenges and opportunities in an electronic environment, and offers some solutions for a comprehensive response to those challenges.

I. What is at Stake?

The paper begins by examining why privacy matters, both in an online and offline environment. It first considers privacy as a matter of individual liberty, autonomy, and even a fundamental human right. All these perspectives remain applicable in a health context, but in addition, breaches of confidentiality are harmful because they can lead to so-called "privacy protective behavior," in which patients avoid seeking health care in order to protect their personal information. Such behavior has a toll on both individual health and, more generally, on public health. It suggests just one important reason why we need to build confidentiality and security into a networked environment.

II. Health Privacy: Definitions and Underlying Concepts

This section considers the concept of privacy, both as it applies to a general environment and more specifically to the medical context. It begins by considering the historical evolution of the term. In 1890, Samuel Warren and Louis Brandeis famously argued that privacy should be defined as "the right to be let alone." Today, definitions tend more closely to resemble Alan Westin's notion of "informational privacy," which suggests that the concept should be understood as an individual's right to control personal information.

Such a definition is particularly important in a global information age, and this section identifies two considerations that are repeatedly voiced regarding the handling of medical data. The first concerns the almost unlimited uses for medical information. Data gathered in a medical context and used for other purposes, it is argued, poses serious privacy risks. The second concern emphasizes the benefits that can be accrued through medical data. This section points to these tremendous benefits, and argues that, while confidentiality of information is essential, patients may miss out on some of the benefits if data controls in the name of confidentiality over-restrict the uses and dissemination of information. The solution is to find a balance between the potential harms and the potential benefits represented by medical

* **Connecting for Health** thanks Stefaan Verhulst, Chief of Research, Markle Foundation, for drafting this paper.

data. That balance can be achieved through a careful deployment of appropriate technologies, combined with strong laws and other forms of confidentiality protection.

III. Health Privacy in a Digital Health Information Networked Environment: What is Different?

This section argues that existing notions of medical privacy are somewhat outdated in a networked health information exchange environment. It discusses six risks increased by such an environment, arguing that these risks require new and innovative solutions. While some of these risks exist in an offline world, they have become more pronounced, in large part due to the scale of data transactions and the relatively greater ease of collecting, linking, and disseminating information over a network, and to a reduced ability to "leave the past behind" and to shield sensitive information. Among the increased risks include:

1. **Commercial misuses of data**, including the use of medical data to deny or restrict insurance coverage; restrict credit or other financial benefits; or in unsolicited marketing;
2. **Government misuses of data**, including secondary use of personal health information by government agencies (for employment and other purposes) and the need to balance national security with health privacy considerations;
3. **Criminal misuses of data**, including fraudulent acts that result in financial or other harm;
4. **Security breaches**, including hacking and other criminal activities that lead to "data leakage";
5. **Data quality issues**, including data corruption and loss; and
6. **Harmful social consequences**, including stigma, exposure, and embarrassment.

IV. Defining a Comprehensive Privacy Architecture: Establishing Trust in the Network

This section defines some principles for responding to the above risks and protecting medical privacy in a networked environment. It begins by discussing existing privacy protection principles adopted in the United States, the Organisation for Economic Co-operation and Development (OECD), and Canada. It then argues for the following nine principles:

1. **Openness and Transparency**
2. **Purpose Specification and Minimization**
3. **Collection Limitation**
4. **Use Limitation**
5. **Individual Participation and Control**
6. **Data Integrity and Quality**
7. **Security Safeguards and Controls**
8. **Accountability and Oversight**
9. **Remedies**

Together, these nine principles amount to a comprehensive privacy protective architecture that can—and should—be applied in a networked environment.

V. Current Laws and Guidelines and How They Integrate an Architectural Approach

This section includes a brief overview of existing privacy protection laws in the United States. It begins by discussing federal protections, and in particular protections built into the Health Insurance Portability and Accountability Act of 1996 (HIPAA). It then discusses the patchwork of state laws, pointing out that these generally fall into three categories: constitutional protections, common law protections, and statutory protections. Finally, it discusses the emergence of, and potential difficulties and opportunities posed by, new community based health networks.

VI. Conclusion

The conclusion offers a summary of the preceding discussion. In particular, it revisits the nine principles and argues that they need to be considered together, as part of an integrated and comprehensive approach to medical privacy.

Introduction and Overview

As we move towards the creation of a health information environment, the potential for privacy intrusions increases, with potentially devastating impact on quality and access to health care. Any up-front planning should take privacy and security into consideration. This paper starts from the belief that it is possible—and necessary—to build privacy into health information technology (HIT) applications so that its benefits can be maximized. It aims to provide background on what is at stake, what has already been achieved in health privacy, what makes the current environment different, and how to provide for a comprehensive response. The paper provides for nine privacy architectural principles that should guide the design of policies, practices, and technologies to protect privacy in a networked environment. In addition it briefly provides an overview of current attempts to address the privacy and security issues within the context of a networked health information environment.

I. What is at Stake?

Individual Liberty and Autonomy: An International Approach

In many countries and treaties, privacy is considered a fundamental right, equivalent to other basic individual liberties such as freedom of speech and thought. Both the United Nations Declaration of Human Rights and the International Covenant on Civil and Political Rights, for example, recognize the right to privacy. In these treaties, privacy is recognized as a form of autonomy, a way to ensure protection from "arbitrary interference"[1] by the state or other entities. In addition, several broad, international principles exist that have been adopted (and adapted) by a variety of countries. For example, as we shall see, in its 1995 Directive on Protection of Personal Data, the Organisation for Economic Co-operation and Development (OECD) led the way in defining several principles for privacy protection. The European Union (EU) and other countries have subsequently adopted these. Interestingly, this directive differs significantly from the US approach in that it takes a broad, omnibus approach to privacy protection rather than the sector and often state specific approaches adopted in the United States.

Understood in this broad way, as a fundamental human right, a violation of privacy can be considered a serious violation of an individual's basic rights, equivalent, perhaps, to imprisonment without trial or the denial of free expression. Naser and Alpert (1999) point out that this violation is particularly serious in a medical context, where patients are often already somewhat helpless and in a position of dependence.[2] They write: "When patients … disclose intimate secrets about themselves they also become more vulnerable. Patients who are ill already have a diminished sense of autonomy" (22). In such instances, robbing individuals of their privacy is tantamount to a serious violation of their individual liberty.

Privacy Protective Behavior in a Medical Context

In addition to a violation of individual rights, the loss of privacy in a medical context has other negative consequences, some of which can be understood as collective harms. Social scientists have frequently established that surveillance, not just in the medical field, but across fields, can have a "chilling effect" on individual behavior (Alpert 2003; Goffman 1966; Westin 1967). In the medical field, this chilling effect can lead to what experts call "privacy protective behavior" (Goldman 1998, 49). Such behavior includes hiding evidence of pre-existing conditions from doctors or insurance companies; paying out-of-pocket for treatment; or simply avoiding treatment altogether.

Goldman, in a paper on the importance of medical privacy, lists four negative consequences of such privacy protective behavior (Goldman 1998, 49):

[1] United Nations, Universal Declaration of Human Rights, Article 12. Available at: http://www.nps.gov/elro/teacher-vk/documents/udhr.htm.

[2] The EU Directives mentioned above similarly treat medical violations of privacy as particularly egregious cases.

(1) The patient may receive poor-quality care, risking undetected and untreated conditions.
(2) The doctor's abilities to diagnose and treat accurately are jeopardized by a lack of complete and reliable information from the patient.
(3) A doctor may skew diagnosis or treatment codes on claim forms, keep separate records for internal uses only, or send incomplete information for claims processing to encourage a patient to communicate more fully.
(4) The integrity of the data flowing out of the doctor's office may be undermined. The information the patient provides, as well as the resulting diagnosis and treatment, may be incomplete, inaccurate, and not fully representative of the patient's care or health status.

Survey Evidence

These negative consequences are not mere hypotheticals. A large number of surveys over the years have consistently shown that the public is concerned about breaches in confidentiality, and that "privacy protective behavior" is a very real phenomenon. For example, as reported by Janlori and Hudson (141), a 2000 survey of Internet users found that 75 percent of respondents were worried that health sites shared information without consent; and that a full 17 percent would not seek health information on the web due to privacy concerns. Another poll, also conducted in 2000, found that 61 percent of Americans felt that "too many people have access to their medical records."[3] Overall, concern about privacy seems to have increased over time: while a Harris Interactive Inc. poll conducted in 1978 found that 64 percent of respondents were concerned about privacy, a similar poll conducted in 1995 by Harris found the number had increased to 82 percent (Goldman 1998, 50).

The surveys also show that such concerns frequently lead to privacy protective behavior. For example, in a survey conducted by the California HealthCare Foundation, more than one out of six adults said they had done something "out of the ordinary" to hide private medical information (Alpert 305). In another survey conducted by Harris in 1993, 11 percent of respondents said they sometimes chose not to file an insurance claim, and 7 percent said they sometimes neglected to seek care in order to avoid damaging their "job prospects or other life opportunities" (Goldman 1998, 50).

Such behaviors do not just cause potential damage to an individual patient's health. They also impose a collective burden, leading to greater costs and public health problems that an already overstretched health system can ill-afford.

II. Health Privacy: Definitions and Underlying Concepts

Understanding the concept of privacy is essential to designing better policies, practices, and technologies to protect consumer and individual privacy. The trouble, however, as one observer points out, is that "privacy is a notoriously vague, ambiguous, and controversial term that embraces a confusing knot of problems, tensions, rights, and duties" (Bennett 1992, 11-12). In attempting to define privacy, one expert resorts to a version of Justice Potter Stewart's famous definition of pornography, arguing simply that: "You know it when you lose it" (Goldman 1999, 101). In an effort to lay the foundations for our following discussion of policies and principles, this section attempts to provide a certain amount of conceptual clarity to the idea of privacy.[4]

Privacy as a General Concept

One of the earliest definitions of privacy was published in 1890, in a *Harvard Law Review* article by Samuel Warren and Louis Brandeis. In that article, entitled "The Right to Privacy," Warren and Brandeis argued that privacy could be defined as "the right to be let alone." The article was drafted in response to concerns over the potential privacy violations that would occur as a result of a new technology. Warren and

[3] These and more survey results can be found at: http://www.epic.org/privacy/survey/.

[4] While much of this discussion refers to broad federal approaches to privacy, it is essential to recognize that privacy protections in the United States have been far more localized and sector-specific. Indeed, the states, not the federal government, have generally led the way in protecting privacy.

Brandeis were writing about the modern press, and particularly the instantaneous photograph, which they felt invaded "the sacred precincts of private and domestic life."

More than 100 years later, we continue to grapple with difficult problems surrounding privacy, and once again, the concern is largely driven by technology. The now-classic definition of privacy in the information age was supplied by Alan Westin, who in his 1967 book, *Privacy and Freedom*, argued that: "Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" (7).

Westin's definition of privacy is probably the most prevalent, and widely-accepted, today. It is sometimes referred to as "informational privacy,"[5] and it is easy to see why this notion of privacy would have particular relevance in the digital era. In 1971, Harvard professor Arthur Miller predicted that all individuals would eventually be the subjects of a "womb-to-tomb dossier." Westin himself argued that, in the information era, every individual was accompanied by a "data shadow" which could reveal even the most intimate and apparently mundane details about his or her life.

Privacy is Not a Static Concept

Such a data shadow, if it indeed materialized, could seriously threaten individual privacy and, by extension, a host of other liberties that citizens in modern Western democracies take for granted. Michael Froomkin, for example, has predicted the "death of privacy." It is important to recognize that the notion of privacy is not static. It changes with time, as the evolution from Warren and Brandeis' concept to Westin's definition makes clear; and it changes depending on the field or environment to which it is applied. This means that privacy is a malleable concept; its treatment and protection can be changed to suit public concerns. In the following sections of this paper, we show how certain protections can be established in response to current concerns over privacy. First, it is important to understand how the concept of privacy is context-sensitive. It is sensitive to particular historical moments. In a more recent article, Westin argues that post-war understandings of privacy have undergone four distinct phases (2003, 434). These include:

- A Privacy Baseline Phase, which ran from 1945 to 1960, and was marked by a relative inattention to, and lack of concern regarding, privacy issues;
- The First Era of Contemporary Privacy Development (1961-1979), which for the first time "marked the rise of information privacy as an explicit social, political, and legal issue";
- The Second Era of Privacy Development (1980-1989), which continued some of the concern begun in the First Era, but overall, "can be seen now as a period of relative calm before the storm";
- And finally, the Third Era of Privacy Development (1990-2002), which "is the period when privacy became a first-level social and political issue in the United States, assumed global proportions, and was impacted by 9/11 and its aftermath."

Privacy and Health

In addition to these well-defined periods, privacy can also be applied to a range of distinct issues; it is sensitive, too, to the field or realm within which it is applied. National security, commerce, and fraud all have privacy dimensions. Although many of these may overlap, there might also be some differences. It is therefore useful to spend some time on the trajectory of privacy as a medical concern. This is particularly important because, as Westin points out, health plays a critical role in his Third Era. Indeed, Westin explicitly points to the rise of genetic testing and the possibility of electronic health records as concerns in this new era (2003, 442).

Although health may have risen to the top of the privacy agenda in recent years, it has long been a topic for privacy advocates and policymakers. As pointed out in a recent report

[5] The US National Information Infrastructure Task Force defines the term as: "Information privacy is an individual's claim to control the terms under which personal information—information identifiable to an individual—is acquired, disclosed, and used." (See http://www.disastercenter.com/Html/PrivacWG.htm; http://www.datenschutz-berlin.de/gesetze/internat/een.htm)

by the Health Privacy Working Group, an initiative, comprised of diverse health care stakeholders (plans, providers, accreditors, and scholars), located at Georgetown University and directed by the Health Privacy Project, national attention to medical privacy can be traced back at least to 1973, when "there were calls for increased attention to the privacy concerns presented by the use of computers in the health care industry" (10).[6] Janlori Goldman also points out that the guidelines and codes of practice developed by the US Department of Health, Education, and Welfare in 1973 continue to serve as the underpinnings for a variety of privacy laws across sectors, suggesting the central role always occupied by concerns over medical privacy (1999, 103). The Privacy Protection Study Commission, created by the Privacy Act, expressed some of those concerns in 1977. "It appears," wrote the Commission,

> *"that the importance of medical-record information to those outside of the medical-care relationship, and their demands for access to it, will continue to grow ... There appears to be no natural limit to the potential uses of medical-record information for purposes quite different from those for which it was originally collected."* [7]

In these and other discussions of health as a privacy concern, at least two distinct themes can be identified. The first, pointed out by Sheri Alpert in a wide-ranging review of the literature on medical privacy, is evident in the above quote, and particularly in the Privacy Protection Study Commission's concern that "there appears to be no natural limit" to the uses of private medical data. As Alpert puts it, there is a recurring concern in the literature over the potential "harm that can befall patients if their medical information is disclosed either in ways that exceed their expectations or if information reaches the hands of people who should not have access to it" (Alpert, 304). She cites a number of authors expressing concern over such potential misuse, and argues that the primary purpose for a patient's personal health information is—and should be—"the clinical diagnosis, treatment, and care of that patient" (305).

The second recurring theme is somewhat contradictory. It provides a counter-argument to Alpert's point, emphasizing the tremendous potential benefits that can be accrued through medical data. Briefly, it is anticipated that the use of medical data, particularly when enabled by electronic health records, has the potential to transform the way patients receive care, and to introduce a far greater degree of efficiency and effectiveness in our nation's medical care system.

Individuals recognize these potential benefits. The same surveys that reveal concern over privacy also show that people are eager to exploit the potential benefits of new technologies. A study conducted by Foundation for Accountability (FACCT) for **Connecting for Health** revealed that while 70 percent believe a personal health record would improve quality of health care, almost all respondents (91%) indicated that they were very concerned about privacy and keeping their health information secure.[8] Likewise, a 2005 survey conducted by the consulting firm Accenture found that an overwhelming number of respondents thought medical care would improve if doctors had access to electronic medical records (EMRs); at the same time, asked to rank their top five concerns with EMRs, respondents put privacy at the top of the list.[9] In recent congressional testimony, Westin stated that "surveys show that most consumers want the opportunities and benefits of our consumer-service and marketing-driven society. With proper notice and choice, more than three out of four consider it acceptable that businesses compile profiles of their interests and communicate offers to them." He pointed out that some 63 percent of Americans, or 125 million people, can be classified as "Privacy Pragmatists": they are willing to share a certain amount of information in the interests of greater efficiency and service,

[6] For a copy of the report, see http://www.healthprivacy.org/usr_doc/33807.pdf.

[7] http://aspe.os.dhhs.gov/datacncl/PHR1.htm.

[8] Available at: http://www.connectingforhealth.org/.

[9] See http://www.accenture.com/xd/xd.asp?it=enweb&xd=_dyn\dynamicpressrelease_857.xml.

as long as they know their information will be safeguarded with privacy protections.[10]

One of the central challenges confronting privacy advocates is to find a balance between these two themes—what Westin, writing on the concept of privacy generally, calls the "distinctive balance between the private sphere and the public order" (2003, 432). Much as it is essential to protect confidentiality of information, so it is essential for our privacy and information laws to maximize the potential benefits that can be offered by medical data. Patients must not feel that their information is misused in any way that violates their privacy; but equally, if information is not shared or disseminated at all, then patients themselves will be the losers.

The solution to achieving this balance lies in well-defined principles that protect information while permitting it to be shared in a meaningful and productive way. Building on the recommendations of the Health Privacy Working Group (many of which were included in HIPAA Privacy Regulations), this backgrounder discusses steps to "integrate privacy protections as part of information practices" (8). This process of integration, in which confidentiality and security protections are built into the architecture of electronic health records and other means of using data, is the best way to ensure that the full benefits of information technology are realized while at the same time protecting the confidentiality and security of personal health information.

III. Health Privacy in a Digital Health Information Networked Environment: What is Different?

We have seen that conventional notions of privacy are today equated with the right to protect information about one's self. The right to privacy may therefore be thought of as a right to secrecy, and privacy protections, whether legal or otherwise, commonly designed to remedy "invasions of secrecy", for example, through illegal entry into an individual's home. Such protections are often designed with reference to an individual's "right to consent" i.e., confidentiality is typically protected by the principle that individuals must give their consent before information about them is allowed to leave the protected domain.

As we shall see, these principles are somewhat outdated in the context of an electronic network. In particular, the widespread availability of databases containing personal information challenges the "right to consent" and "invasion" principles upon which so many privacy protections are currently based. For example, when an individual's personal health information is aggregated with other patients' data and resold as part of a database, no opportunity is given to the individual in question to provide consent on reuse of that information. Indeed, in many cases an individual will not even know that his or her personal health information has been reused.

The new environment poses a host of additional challenges to existing privacy protections and principles. If we are to develop effective solutions, it is essential to better understand these new challenges. It needs to be clear, at the outset, that while a digital and networked environment offers much potential and many opportunities, it also poses several new categories of risk. This section will explore some of those risks.

After exploring those new risks, this section will discuss some privacy architectural principles to deal with those risks. A central principle of this backgrounder is that new privacy challenges cannot be addressed solely by focusing on post-violation remedies and penalties, but also (and more importantly) through network architectures that govern the information flows and the handling of personal information. Such architectures must be designed in a way to protect privacy before violations occur. Therefore, after outlining the new risks, we argue that privacy in a digital setting requires *structural and systemic approaches*.

New Environment, New Risks

1. Commercial Misuses of Data

Perhaps the most serious—and probably pervasive—privacy violations in the information age stem from the potential for commercial

[10] http://energycommerce.house.gov/107/hearings/05082001Hearing209/Westin309.htm.

misuse of data. In recent years, an extensive data market has developed, driven largely by data aggregators or "data brokers." These data brokers collect, repackage, and sell information that is either available in the public domain, or they illicitly aggregate data that was collected for another purpose from that for which it is ultimately used.[11] Deborah Platt Majoras, Chairman of the Federal Trade Commission, described the general data market of personal information in recent Senate testimony:

> The information industry is large and complex and includes companies of all sizes. Some collect information from original sources, others resell data collected by others, and many do both. Some provide information only to government agencies or large companies, while others sell information to small companies or the general public.[12]

The emergence of data as a commodity, traded in often-opaque information markets, has led to serious concerns about privacy. In *No Place to Hide,* Robert O'Harrow describes in vivid detail the wealth of information that now exists on individuals, and the various and frequently harmful ways in which it can be used, often without the individual's knowledge or consent. Some possible harms include:

- *Denial or Restrictions on Insurance Coverage and Other Benefits:* Information acquired in one medical setting (for example, routine testing) can become part of an individual's data shadow and later be acquired by insurance companies to deny or otherwise restrict coverage. At least two companies, the Medical Information Bureau (MIB) and AllClaims, currently offer databases of patient medical records to insurance companies and others. Such commercial use of data represents a serious problem in part because inadequate insurance cover creates new and potentially serious public health problems. It is also a serious concern because, as we have seen, knowledge of such risks leads to privacy protective behavior on the part of individuals that can pose further health problems.

- *Restrictions by Credit or Other Agencies:* Medical data can also be acquired commercially and used by non-medical agencies like credit card companies or banks. If such data points, for example, to a serious underlying medical condition, it can lead to denial of credit, mortgages, or other financial services. This "leakage" of information from a medical to non-medical setting is a serious problem in an era of data aggregation; it shows the need not only to build privacy protections within the health care sector, but also to develop strict procedures to control transmission of information across sectors.

- *Unsolicited Marketing:* Data acquired commercially can also be used by pharmaceutical companies and others to market drugs based on information about individuals' medical condition. Two notable cases of such marketing occurred in 1998, when CVS and Giant Food, two pharmacy chains, offered patient prescription records to private companies that later used the records as the basis for marketing. In addition to the underlying privacy violation involved in making the data available, it can also be argued that the unsolicited marketing itself represents a privacy violation.[13]

2. Government Misuses of Data

The debate over privacy and data aggregation often refers to commercial uses of data. However, the state also makes frequent use of an individual's medical data shadow for law

[11] The illicit use of data is not particular to the networked environment. What has changed is the scope of potential violations: As the network expands and as the amount of data increases, so does the possibility of confidentiality violations. In addition, a networked environment facilitates the illicit acquisition (e.g., through theft) and dissemination of data. This is in large part due to digitalization of information, which is easier to store, and to steal without its original owner even noticing.

[12] http://judiciary.senate.gov/testimony.cfm?id=1437&wit_id=4161.

[13] Goldman (1998, 10).

enforcement and other purposes. In 1998, for example, police in Virginia, investigating a car theft from a parking garage near a drug treatment center, collected 200 medical records as part of their investigation; they later acknowledged their actions as an unnecessary violation of patient privacy. State welfare agencies and the Immigration & Naturalization Service (INS) have also used welfare and immigrant health records in the administration of their respective programs.[14]

An emerging category of risk that is particularly worth highlighting stems from the increasing capability of governments to indulge in surveillance activities. A recent report, jointly issued by the American Civil Liberties Union (ACLU), Focus on the Global South, Friends Committee (US), International Civil Liberties Monitoring Group (Canada), and Statewatch, highlights the risk.[15] It argues that individual pieces of information on travel and other practices that are currently being collected could lead to an international surveillance framework that "dwarfs any previous system and makes Orwell's book *Nineteen Eighty-Four* look quaint." These individual pieces include registration of foreigners, national ID policies, and biometric identification methods.

The report also points out that much of this information is collected in the name of national security. The authors argue that the information will not fulfill its stated purpose, but the stated reason for collection does point to a complication in addressing privacy violations by the state, namely, that government collection and use of data often has legitimate and vital national security purposes. In a post-9/11 environment, in particular, data can be useful in stopping terrorist attacks before they occur. A national information network is today considered critical to enhancing the nation's intelligence programs. As many—including the Markle Foundation—have argued, however, it is essential that such a network be designed with built-in protections for privacy.

Such protections would be both architectural (i.e., built into the design of the network), practices, *and* policy-based. We discuss architectural solutions below. One important policy step involves reform of the 1974 Privacy Act. In recent Senate testimony, James Dempsey, the Executive Director of the Center for Democracy & Technology (CDT), pointed out that government use of data is susceptible to privacy violations due to shortcomings in that act, which requires government agencies to collect and use data subject to the provisions of the Fair Information Practices. But as Dempsey further pointed out, such protections are only relevant to "federal 'systems of records', [meaning] ... that the government can bypass the Privacy Act by accessing existing private sector databases, rather than collecting the information itself." He went on to describe the possible negative consequences that can occur when the government accesses private data without the restrictions of the Fair Information Practices:

> *[A]lthough the Privacy Act requires notice to and consent from individuals when the government collects and shares information about them, gives citizens the right to see whatever information the government has about them, and holds government databases to certain accuracy standards, none of those rules applies when the government accesses commercial information without pulling that data into a government database. Currently, the government need not ensure (or even evaluate) the accuracy of the data; it need not allow individuals to review and correct the data; and the government is not limited in how it interprets or characterizes the data.*[16]

3. Criminal Misuses of Data

Both commercial and government uses of data have legitimate purposes; generally, misuses and privacy violations represent the exception rather than the norm. But digital data, medical or otherwise, is also susceptible to criminal misuse, which can result in serious violations of privacy, considerable financial expense, and even physical injury and death.

[14] Health Privacy Working Group (1999, 10).

[15] http://www.theregister.co.uk/2005/04/21/icam_surveillance_report/.

[16] http://judiciary.senate.gov/testimony.cfm?id=1437&wit_id=2875.

Identify theft, in which criminals acquire Social Security numbers or other identifying information, represents a particularly serious problem. In 2003, the Federal Trade Commission (FTC) estimated that 10 million Americans (nearly 5 percent of the adult population) were victims of some form of identity theft.[17] According to the FBI, the Internet Crime Complaint Center (IC3), a joint project between the FBI and the National White Collar Crime Center, received more than 100,000 complaints regarding identity theft in the 5-year period between its opening in 2000 and 2005. It estimated the costs of identity theft as nearly $40 billion annually, not including credit card fraud.[18]

For all its seriousness, identity theft represents just one possible instance of criminal misuse of data. It imposes substantial financial costs, but other types of illegal activity can result in even more dangerous consequences. Consider the following two examples:

- In 1999, a woman named Amy Boyer was murdered as the direct result of her data shadow. She was killed when a man purchased her Social Security number, address, and other information from a data broker called Docusearch (the man paid just $154). The information was used by the man, who had been obsessed with Boyer since her youth, to find her place of work and kill her.[19]
- Concerns about similar criminal misuse of data were also raised in a 2005 case brought by a Juneau, Alaska nurse who sought to have her address removed from public records, a licensing condition for all nurses. Expressing a fear of stalkers, she argued, with the assistance of the ACLU, that making her address publicly available posed a serious threat not only to her privacy, but also to her physical safety.[20]

[17] http://judiciary.senate.gov/testimony.cfm?id=1437&wit_id=4161.

[18] http://judiciary.senate.gov/testimony.cfm?id=1437&wit_id=4162.

[19] http://judiciary.senate.gov/member_statement.cfm?id=1437&wit_id=2629.

[20] http://www.adn.com/news/alaska/story/6399520p-6278454c.html.

4. Security Breaches

As the above examples illustrate, data can be acquired and misused by criminals in two ways:

- Through legal means, by following or purchasing a legitimate data trail. In such cases, it is the subsequent misuse that is illegal, not the acquisition of the data itself.
- Through criminal acquisition and use of data, in which the way the data is collected is itself illegal. Such criminal acquisition frequently arises as a result of security breaches, discussed in this section.

Security breaches, sometimes referred to as "data leakage," represent a serious category of risk in the information age.[21] They are not unique to the information age, but digital records and networks present particular vulnerabilities that do not exist in a paper-based world. These risks include the relatively greater ease of remotely hacking a network than physically breaking into a paper records depot; and the fact that large quantities of data are stored on servers and hard disks that are connected to the world, protected only by firewalls or other imperfect security protocols. In addition, digital data is much easier to replicate, and such replication can be done without damaging or removing the original, making it easier to acquire data illegally without the owner even being aware.

These and other factors make it easy to steal or criminally acquire data in the information age. Recent examples suggest that criminals are well aware of network vulnerabilities and that criminal acquisition of data is a growing risk. Recently, for instance, Ameritrade, an online broker, announced that it had lost a tape backup containing data on 200,000 current and former customers. This followed announcements by Lexis Nexis that up to 310,000 customer records may have been hacked; and reports by ChoicePoint, a data aggregator, of similar violations.

Such examples highlight the inherent vulnerabilities of networks and information stored in a digital format. While we have

[21] For a listing of recent security breaches and data violations, see http://www.privacyrights.org/ar/ChronDataBreaches.htm.

outlined some of the security vulnerabilities, many more exist. Of course it is impossible to fully protect a network against all forms of intrusion—the best we can hope for is to minimize intrusions.[22] The important point is that the existence of such vulnerabilities requires architectural solutions that build security protections from the start, rather than post-fact remedies. We discuss some possible architectural solutions in the following section.

5. Data Quality Issues

In addition to introducing a greater potential for security breaches, a digital environment also introduces potential data quality issues. Problems with data quality, which include data loss or corruption, are not traditionally thought of as privacy violations, but are closely interrelated with current privacy concerns.

Consider, for example, some recent anecdotes regarding the wrongful inclusion of individuals on national no-fly lists or other terror databases. Inclusion in such databases can be considered a privacy violation on at least two counts. First, it can automatically lead to private data being viewed by a range of agencies and groups, which could claim access on national security grounds. For example, if an individual is wrongly placed on a federal no-fly list, local law-enforcement agencies might also gain access to that individual's information based on law-enforcement sharing procedures.

Second, and more relevant to a discussion of medical privacy, it is important to recognize that much as individuals can be placed by mistake on no-fly lists, so they can be included in medical databases with false identifying information. Patients could, for example, be denied insurance based on mistaken information regarding medical conditions; similarly, they could be forced to pay higher life insurance or other premia.

It is important to acknowledge that, for the moment, such risks remain often theoretical, and that they are not particular to the online world, but also exist in a paper-based system of records. Nonetheless, they highlight the need not only to build strong privacy protections into network architecture, but also remedies and means of appeal against data quality issues. If patients are not able to have privacy or data quality grievances addressed in a quick and clearly identifiable manner, there is a danger that those grievances will be compounded. In addition, a comprehensive approach to data quality must include procedures to ensure information integrity to prevent errors from occurring in the first place.

6. Harmful Social Consequences

Finally, while much analysis of privacy focuses on adverse economic or health consequences, it is important to recognize that privacy violations can impose a very real social cost on individuals, making it difficult for them to live meaningful lives within their communities. One notable example occurred in 1998, when a San Diego pharmacist revealed a man's HIV-positive condition to his ex-wife. The man, who was locked in a custody battle with the woman in question, ultimately settled the case rather than face the stigma of his condition being made public.[23]

The need to carefully control such social consequences is all the more apparent when we consider that societies also use such "shaming" techniques as regular tools for law-enforcement procedures. Consider the widespread use of so-called Megan's Laws to maintain public sex offender registries. The use of such legitimate and legal shaming techniques makes it essential to draw up strict rules to differentiate between acceptable disclosures of personal information in the public domain, and unacceptable disclosures.

Writing more than 200 years ago, Adam Smith, often considered the father of modern economics, argued that material well-being was just as important to human happiness as "the right to appear in public without shame." This argument is as true today as it was then, and it draws attention to the very real need for controls on how information about an individual is released into the public domain, and shared with a community.

[22] See for instance Paul Clayton (Chair): *For the Record: Protecting Electronic Information*; National Academy Press, 1997.

[23] Health Privacy Working Group (1999, 10).

IV. Defining a Comprehensive Privacy Architecture: Establishing Trust in the Network

The previous section described some of the categories of risk represented by new technologies and methods of information dissemination. Clearly, these risks and vulnerabilities require new responses. These responses, moreover, must not be ad-hoc or post-fact, but designed in a systematic and comprehensive manner. At the core of adequate privacy protection in the digital age is that it must be supported by policy, practice, and the architecture of the network.

The purpose of this section is to provide privacy architectural principles for the policy, technology and, more generally, for the social and economic context within which the technology is used. In what follows, we present nine core principles of privacy protection based upon Fair Information Practice Principles (FIPPs) and explain how they must be built into the way information is collected and shared. Before that, we review currently existing Fair Information Practice Principles.

Throughout this discussion, we must keep in mind that to be effective, the scope of the protection will need to be determined and defined. This requires considering whether different kinds of protections should apply for different kinds of data; the kind of relationship and the level of trust (either socially, contractually, or legally determined) one aims to address and achieve. In addition, one needs to focus on the various systems of records or the information flow and any third party that maintains those systems.

Fair Information Practice Principles

Before discussing our core principles for a networked environment, it may be useful to briefly consider some existing principles for privacy protection. These principles provide a useful template, but they are not optimized for a network-driven world. Many were designed long before the age of the Internet, data brokers, and data aggregation. As such, they may need to be tailored, adapted, and, in some cases, expanded to address the specific risk management challenges posed by the digital age in general, and the rise of EMRs in particular.

The Privacy Rights Clearinghouse, a nonprofit consumer group located in California, provides a useful review of existing Fair Information Practices.[24] Here, we provide a summary, based on that review, of existing privacy laws in three jurisdictions:

1. The ***United States***, including the 1973 Fair Information Principles and the 1974 Privacy Act;
2. The ***OECD***, including the 1980 Guidelines on the Protection of Privacy and Transborder Flows of Personal Data; and
3. ***Canada***, including the 1995 Canadian Standards Association Model Code for the Protection of Personal Information.

1. The United States

The Fair Information Practices were implemented over thirty years ago (1973), when the US Department of Health Education and Welfare (HEW) formed a task force to consider the privacy effects of the spread of computer medical records. The Code of Fair Information Practices developed by this task force includes the following principles:[25]

1. **Collection limitation:** There must be no personal data record keeping systems whose very existence is secret.
2. **Disclosure:** There must be a way for individuals to find out what information about them is in a record and how it is used.
3. **Secondary usage:** There must be a way for individuals to prevent information about them that was obtained for one purpose from being used or made available for other purposes without their consent.
4. **Record correction:** There must be a way for individuals to correct or amend a record of identifiable information about them.
5. **Security:** Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the

[24] See http://www.privacyrights.org/ar/fairinfo.htm for a full discussion.

[25] Reproduced from "The Law of Privacy in a Nutshell," Robert Ellis Smith, *Privacy Journal*, 1993, pp. 50-51.