Regulatory Compliance and the Correlation to Privacy Protection in Healthcare

Tyrone Grandison[+], Rafae Bhatti[*]
[+]IBM Almaden Research Center,
650 Harry Road, San Jose,
California 95120, USA
[*]Oracle Corporation
500 Oracle Parkway, Redwood Shores,
California 94065, USA

Abstract. Recent government-led efforts and industry-sponsored privacy initiatives in the healthcare sector have received heightened publicity. The current set of privacy laws and regulations mandate that all parties involved in the delivery of care specify and publish privacy policies regarding the use and disclosure of personal health information. Our study of actual privacy policies in the healthcare industry indicates that the vague representations in published privacy policies are not strongly correlated with adequate privacy protection for the patient. This phenomenon is not due to a lack of available technology to enforce privacy policies, but rather to the will of the healthcare entities to enforce strong privacy protections and their interpretation of minimum compliance obligations. Using available information systems and data mining techniques, we describe an infrastructure for privacy protection based on the idea of policy refinement to allow the transition from the current state of perceived to be privacy-preserving systems to actually privacy-preserving systems.

Key words: Healthcare, Privacy, Regulation, Compliance, Policy, Refinement

INTRODUCTION

In the healthcare industry, privacy concerns are among the main inhibitors to the deployment and use of electronic records systems. In the last decade, the increase in the number of data breaches (Privacy Rights Clearinghouse, 2009) has led to an increase in the number of companies who are concerned about data and brand protection, which has translated into increased spending on healthcare privacy compliance efforts. In the United States, the Health Insurance Portability and Accountability Act[1] (U.S. Department of Health and Human Services, 1996), the new security and privacy requirements imposed by the Health Information Technology for Economic and Clinical Health Act[2] (U.S., 2009) and the changes to HIPAA mandated by the American Reinvestment and Recovery Act[3] (U.S.A., 2009) are normally assumed to provide the baseline for privacy compliance for healthcare entities.

As healthcare organizations implement the required privacy policies, what remains to be ascertained is the impact these policies have on the improvement of privacy practices. More specifically, we address the question: *How well does the use of privacy policies translate into good privacy practices?* The *use of privacy policies* refers to the specification, notification and enforcement of policy; while *privacy practices* refer to the processes and mechanisms (i.e. technological and otherwise) that enable the safe handling of sensitive information.

The answer to our question lies in the design and enforceability of the policy itself. As we will reveal, a policy may be designed to cover all the relevant provisions of the regulation, and yet may still be vague enough to afford very little privacy protection to the patient. We will discuss this further in a later section. Concerns about the inadequate state of privacy protection despite the enactment of data protection regulations have long existed in mainstream media (Pear, 2009). In addition to design issues, studies also indicate that the enforcement of policies governing the use of protected patient information in current healthcare information systems is also lax and that policy is often bypassed or subverted during regular operation (Rostad & Edsburg, 2006).

This scenario makes it possible to purport compliance with privacy regulations, while engendering a false sense of security (or more aptly, a false sense of privacy) among patients. It makes the existence of a policy, in the first place, insignificant; as it does not precisely represent the company's true stance on data protection. Also, this undermines the notion of empowering the patient, as his consent to a policy is no longer a genuine reflection of the company's privacy practices. In an electronic health records environment, this conundrum highlights the need for privacy enhancing technology. No prior work has investigated how stated privacy policies measure up to the levels of protection required to truly ensure the safety of patient data, and whether the current system can be elevated from one that purports regulatory compliance to one that really safeguards the privacy of healthcare data. Our goal is to contribute to the solution of this pressing need.

We believe that it is possible, and desirable, to define appropriate mechanisms to ensure that privacy protection moves from the adherence to minimum standards to a level that truly reflects good privacy protection for patients. In this paper, we first evaluate the current *HIPPA-inspired* privacy practices against the needs of the patient and then present a privacy management architecture called PRIMA that enables refinement of privacy policies based on actual practices of the organization. *Policy refinement* helps mitigate the stated conundrum because it allows one to (i) improve the design of the policies in order to elevate the level of privacy protection afforded to the patient, and (ii) better align the policies with actual privacy practices of the organization.

The rest of the paper is organized as follows. First, we summarize the base constructs for a discussion on healthcare privacy in the United States. Then, we describe our survey of actual privacy policies used by healthcare organizations and assess them. This is followed by a description of infrastructure for privacy protection that is based on this notion of *policy refinement*, which enables improved privacy practices in healthcare. Finally, we conclude with a synopsis of the insight gained from this work.

BACKGROUND

Several data protection regulations have surfaced around the globe (Wong, 2009) over the last few decades that directly impact healthcare privacy. Among these are the Personal Information Protection and Electronic Documents Act[4] (Office of the Privacy Commissioner of Canada, 2000), the Personal Data Protection Law (Japanese Ministry of Internal Affairs, Communications Information, and Communications Policy, 2003), and laws enacted pursuant to the European Union Directive on Data Protection. In the United States, a combination of federal laws, state laws, and common law, i.e. tort and contract, requirements define the bounds of privacy protection. In the past few years, HIPAA compliance has become the measure of adequate privacy protection for the healthcare sector. This is becoming more relevant with the recent reports encouraging a transition to a secure, private, interoperable electronic health infrastructure (U.S., 2009; HIMSS, 2009; U.S. PITAC, 2004).

To ground our discussion, the rest of this section will focus specifically on privacy protection in healthcare sector as mandated by the HIPAA Privacy Rule (U.S. Department of Health and Human Services, 1996), and the amendments and additions mandated by HITECH (U.S., 2009) and ARRA (U.S.A., 2009).

The Legal Requirements

As a pre-requisite to summarizing the requirements of the HIPAA Privacy Rule, we must define fundamental terms used. *Covered entities* refer to health plans, health care providers and health care clearinghouses. *Protected Health Information (PHI)* refers to all individually identifiable health information held or transmitted by a covered entity or its business associate, electronically, on paper, or orally. Appendix A presents the allowable disclosures under HIPAA and highlights in greater detail the provisions of the HIPAA Privacy Rule. Here, we present the high-level overviews of the five key principles of the HIPAA Privacy Rule assertions:

1. **Notification** - Patients should receive a notice of a covered entity's privacy practices.
2. **Authorization and Consent** - Written authorization is required for disclosures not permitted under the Privacy Rule.
3. **Limited Use and Disclosure** - Covered entities must use or disclose the *minimum necessary* PHI for a specific purpose and ensure the development and implementation of policies and procedures governing access and use.
4. **Auditing and Accounting** - Patients have the right to an accounting of all disclosures of their PHI for non-allowed HIPAA operations.
5. **Access** - Patients have the right, under most circumstances, to access the covered entity's designated record set. Covered entities must amend information that is inaccurate or incomplete.

The HITECH Act (U.S., 2009) augmented HIPAA by 1) strengthening the enforcement risk, e.g. increasing penalties for HIPAA violations, 2) creating a breach notification requirement for the healthcare industry, 3) extending the coverage of the HIPAA requirements to business associates, 4) re-enforcing the audit and authorization rights of patients, and 5) prohibiting the sale of an individual's health information without their authorization.

In addition to solidifying provisions in HITECH, ARRA (U.S.A., 2009) further clarified and refined issues that were previously deemed nebulous, e.g. restrictions on healthcare information sharing for self-pay scenarios, limited data sets, minimum necessary, marketing provisions, etc.

As both HITECH (U.S., 2009) and ARRA (U.S.A., 2009) are fairly recent and most organizations have not had time to updated their policies to reflect the new mandates, in order to have a fair and standard platform our analysis is based on the original HIPAA Privacy Rule mandates. For the rest of this text, reference to *policy* is to a *technical privacy policy artifact* and not to a *legislative policy artifact*. However, the assumption is made that technical policy should be a mechanism to support and instantiate legislative policy.

The first step in enabling privacy compliance, under HIPAA, is defining the entity's privacy policy rules such as which users may use or disclose data, for what purposes will it be used, under what conditions and who will be the end consumers of this data (i.e. the recipients).

P3P and Privacy Policies

Platform for Privacy Preference (P3P) (Cranor et. al., 2002) defines a standard XML format for a computer-readable privacy policy called a P3P policy. A P3P policy includes elements that describe the kinds of data a web site collects, the purposes for which data is used, potential data recipients, data retention policies, information on resolving privacy-related disputes, an indication as to whether a site allows individuals to gain access to their own data, and other information (Byers et. al., 2003). The purposes for which data may be used are typically divided into categories such as current (for which information is being directly supplied), develop (site administration), admin (website improvement), pseudo-analysis or individual analysis (user profiling and customization), telemarketing, and users are provided with the option to consent to the disclosure of their personal information for an indicated purpose. User can supply their preferences using the P3P Preference Exchange Language[5] (Byers et. al., 2003). In other words, privacy preferences encoded as rules in APPEL can be used to evaluate a P3P policy against the preferences of the user. Many tools, such as the Privacy Bird engine (Byers et. al., 2003), exist to allow such an analysis.

The adoption and significance attributed to P3P varies across the market segments and many researchers have also pointed out weaknesses in, and extensions to, P3P (Karjoth et. al., 2002; Schunter et. al., 2002; Li et. al., 2003). However, P3P is now considered a well-accepted and well-adopted standard in electronic commerce (Byers et. al., 2003) and is used by a majority of websites to make customers aware of the policy of a company regarding the protection of their privacy.

HEALTHCARE PRIVACY POLICIES SURVEY

In order to answer our initial question, we commissioned a study of patients, gave them a survey based on the five HIPAA Privacy principles articulated above and collated their expectations with regards to each category. We then conducted a survey of the healthcare privacy policies, which were selected from twenty healthcare companies on Thomson Reuters Top 100 Hospitals list. For each policy, we analyzed it against both the regulation and patient expectations.

Before we proceed, we must further clarify our use of the term *privacy policy*. A typical covered entity in our survey provided an electronic copy of its *HIPAA Notice of Privacy Practices*, as required by regulation. This document specifies how the organization maintains the privacy of members' medical information. An electronic copy of this notice is posted on their websites. Most organizations in our survey also posted a separate *Website Privacy Policy*, which only applies to information collected, used and or disclosed through the company's website. For our purposes, we use the term *policy* or *privacy policy* to mean the virtual combination of both. This ensures that our survey results are valid for the entire electronic healthcare experience.

Generally, the surveyed companies structured their policies to convey information on the following areas: *Collection of Information*, *Information Types*, *Information Use* and *Changes to Information*. There were slight differences in terminology amongst the policies, but the higher level concepts were equivalent. We found the policies to be very clear in articulating the information that will be collected. This information falls into one of three classes: 1) *Protected Health Information*, which includes name, address, social security number, email address, licensure, certifications, education and employment history, etc. and is normally assumed critical for the delivery of care and the company's normal business functions, 2) *Derived Information*, which includes individual access history and usage patterns, which is gathered through *cookies* in order to improve their site and allow personalization or customization, and 3) *Aggregate Information*, which is statistical information, consolidated from IP addresses, computer information and locations (amongst other things), for promotion and marketing.

In the following sections, we examine the trends in the policy statements made by the organizations in our survey, in the light of HIPAA privacy provisions.

Notification, Authorization and Consent

At the start of their policies, the healthcare companies either stated (1) that they do not collect personal information from web page visitors, but do collect web usage statistics in the aggregate form and if one wishes to register with them, then personal information will be collected or (2) that by accessing the companies' web pages you have consented to their privacy policy. Both cases lead to a situation where the patient is assumed to have implicitly consented to the privacy policy through the action of browsing the companies' web pages. We note that this does not satisfy the notification requirement in the HIPAA Privacy Rule necessary to use or disclose PHI.

We also noticed that none of the websites actually published a policy in P3P, or any similar privacy language, and only the natural language version is available online for manual review. The fact that no P3P policy is available on the Website precludes us from performing automated interpretation and analysis. Further, the intentional ambiguity in the regulation and its natural language representation mean that they cannot be directly translated in a machine readable form, like P3P. In this regard, healthcare is behind other sectors (e.g. online retail), where posting P3P policies on their website is now common practice. Admittedly, the privacy requirements in healthcare are more complex.

Another particularly alarming trend that was observed in our survey related to the communication of policy updates to the patients. From their policies, fifteen of the twenty of the organizations were content with simply updating the policy on the website, and making it the responsibility of the user to check for policy changes. It is a general theme that privacy policy changes are communicated with very little concern for the patient. The five companies that stated otherwise indicated they would alert the patient (via post, email, etc.) in case of a policy update. Again, we note that this type of notification is generally insufficient to revise policies for previously collected PHI under HIPAA.

Analysis

Current practices around issuing a notice and obtaining consent are not sufficient unless they provide the patient an opportunity to clearly and easily understand the policy and negotiate any objectionable provisions. This will continue to be a manually intensive task unless the policies are presented to the patient in a format that not only highlights the key segments in the policy, but also allows reasonable modifications to be made by the patient at his/her discretion. The use of P3P and APPEL technology, for example, may facilitate this task. Though recent studies have shown that privacy policies are unreadable by their target audience, irrespective of their format, (McDonald et. al., 2009) and that less than 26% of Internet users read privacy policies (Jensen et. al., 2005), we assert that the codification of policy would enable computer to analyze them and visualize potential problems, perhaps based on a specification of the user's concerns or *hot buttons*.

Unfortunately, HIPAA does not require that the notice be issued in P3P or similar machine-readable format that would facilitate automatic interpretation or analysis. It only goes as far as saying that the notice should be placed in "a clear and prominent location" (U.S. Department of Health and Human Services, 1996, pp. 89), which is completely consistent with the regulation's goal of *technology neutrality*. However, more positive impact would be achieved if some guidance was provided with respect to using the most current and expedient communication mechanisms, e.g. emails and instant messaging. Overall, we observed that industry practices regarding privacy policy notification and changes fell short of the spirit of the HIPAA requirements by failing to obtain patient acknowledgments.

Limited Use and Disclosure

On all websites surveyed, use and disclosure of information are associated with a purpose and specific purposes are defined for information. However, we found that all the organizations have defined very broad and all encompassing purposes, which may be used to exploit exceptions in the HIPAA privacy rules. For example, all the policies mention collecting information for the purpose of *administering healthcare*. They settled for a granularity so coarse that it could subsume a huge category of uses and disclosures of information. As a result, a whole host of activities, which the patient may not be in agreement with, could be interpreted as included within these purposes.

For disclosures to third parties and affiliates, it is common to see the phrase "we require the third parties to comply with policy" in privacy policies. However, there are two significant hurdles here. Firstly, a proposition to either *comply with policy* or to have *use limited by policy* is only meaningful if the policy is not broadly defined and implicitly inclusive of a wide range of business functions. Secondly, apart from *business associate contracts* with the third parties that perform services, requiring PHI, for them, there are no guarantees of the actual enforcement of policy on the third

party. Ideally, covered entities will proactively monitor these third parties to assure that they comply with the business associate agreements. However, the policies make no mention of the (general) terms and ramifications of such agreements.

None of the privacy policies surveyed provides a fine-grained list of roles or employee categories who have the authorizations to view specific categories of patient data. For internal use, the collected information is available to all *members of medical staff*. This is the only requirement for being an *authorized* employee. Nowhere are the precise conditions for being *authorized* stated, nor is there any criteria specified under which any exception-based accesses may be granted (such as in *break the glass* scenarios). Overall, the counsel or consent of the patient is not incorporated in assigning more specific access privileges to employees.

Analysis

While HIPAA requires organizations to obtain unambiguous authorization of the patient before use or disclosure of information for a purpose other than what it was collected for, and recommends adoption of the principle of *minimum necessary* disclosure, there exists exceptions in HIPAA to allow organizations to design policies with broadly defined purposes, and still remain regulatory compliant. This concern has also been highlighted in the public media (Pear, 2009). For instance, while *marketing* is identified as a purpose that requires authorization, various sub-categories are defined, such as *communications for treatment of patient*, that are exempt from the rule, making it possible to disclose patient data for marketing under the assumed purpose. Therefore, it may be assumed that the levels of disclosure post-HIPAA will not necessarily shrink, and in fact a data disclosure previously considered a breach may now fall within the folds of the policy to which the patient has consented. Anton et. al. (2007) observe a similar phenomenon, where disclosures increased post-HIPAA.

Although HIPAA requires an organization to develop and implement internal policies and procedures for controlled access to patient information, this requirement is not re-enforced by requiring the use of stringent technical mechanisms. In other words, this requirement may simply be fulfilled by organizations using the minimal set of access controls, and employee training programs. For instance, while researchers (Anderson, 1996) and medical professionals (IHE, 2006) have clearly articulated the need for fine grained access control for patient health records, all organizations in our survey are legally compliant even though they simply chose to give an umbrella authorization to all of their employees under the cover of *members of medical staff*. While an argument can be made for unhindered access to electronic healthcare resources so as to not impede the clinical workflow, the ethically responsible organizations should still do their level best to adopt the principle of *minimum necessary* use and (voluntarily) implement access controls at a finer granularity. We note that in the multi-vendor ecosystem of the American healthcare, most of the vendors have heavy investment in their existing tools and the economic motivation to move to these fine-grain systems may be low.

Additionally, even within organizations that implement finer-grained access control through the use of mechanisms such as role-based access control (RBAC), an over-use of exception-based accesses has been reported in a recent study (Rostad & Edsburg, 2006). While such accesses are usually meant to be exercised in emergencies, the study revealed that they are increasingly being utilized in non-emergency tasks. This effectively means that even in the presence of fine-grained access controls, the frequency with which exceptions (i.e. bypassing of the access control) are utilized can effectively render that system equivalent to one with umbrella authorizations.

Audit and Accounting

The privacy policies of all organizations advise that patients can obtain audit records for information disclosures. All policies mention that protected health information may be disclosed to government and regulatory authorities for compliance with law. Although not explicitly stated on the websites, the literature from the medical community (Anderson, 1996; Blobel, 2004) suggests that most organizations advocate the use of audit trails of all actions pertaining to patient medical records to meet the audit reporting and accounting requirement. Our experience with clients indicates that audit trails do not record all the necessary context information, such as purpose and recipient amongst other attributes. More alarmingly was the tendency of corporate executives to turn off audit systems because of the storage and performance burden incurred when they ran.

Analysis

The fact that current healthcare audit systems do not capture the required context information in order to provide an accounting means that they do not currently meet this requirement. Even though HIPAA requires organizations to account for all activity (including data disclosures) and provide detailed reporting for audit purposes, fulfilling this

requirement by itself would still not be effective in improving levels of privacy protection unless measures are taken to compensate for shortcomings in the data disclosure and access rules in the privacy policy.

When the purpose or authorization is not established at a fine granularity before any disclosure or access is allowed, the burden falls on the audit mechanism to be able to capture any action that may actually constitute as a violation of the policy. Additionally, when an exception-based mechanism is in place that allows users to override normal access controls, the need for audit-based controls is further accentuated.

While an argument can be made that the deterrent factor of audits is more suited to the healthcare sector because of the critical nature of the services provided, it should certainly not become an excuse for failing to do better.

Access

The privacy policies posted on all the websites in our survey indicated that patients have the right to access or update their personal information maintained by the company through phone or email or an online account.

Analysis

Meeting this requirement may not translate to adequate privacy protection for patient. There are several reasons for this. First, the ability of a patient to access or update personal information maintained by the organization provides no measure of how much information is actually protected unless the patient is also in control of the use and disclosure rules, and, based on our preceding discussion, this is not the case. Second, navigating the processes of information access and update can be simple or laborious for the patient depending on the organization. As per the access policy of all organizations, only personal information, such as name and address, can be accessed and updated online. To access or change medical data on the patient maintained by the organization, a written request is required and it can take up to 60 days to receive a paper copy of one's medical record.

Further Observations

From our analysis, the language used in the privacy policies appears to be unnecessarily convoluted. This is corroborated by other researchers in the field (Hochhauser, 2001; Graber et. al., 2002) for healthcare and finance. Given this, it will likely not be understandable by the average patient. Also, the language used in all these policies was clearly ambiguous. For example, one policy in the study states *"……will not sell, license or transmit to anyone any personal information that members or practitioners provide to us online. We may disclose information obtained online to our partners involved in administering or providing services for our health benefits plans*". These are possibly two seemingly contradictory, yet consistent statements that seem to have a nullifying effect on each other. This observation is supported by work in the medical informatics community (Ball et. al., 2007) and the computer science community (Anton et. al., 2007). Additionally, the policies downplayed the privacy risk involved (Pollach, 2007).

Secondly, the use of P3P was restricted only to cookies and exists in their abbreviated form. This raises an interesting issue. In addition to the information explicitly identified as PHI, three companies also classified cookies as PHI, whereas the remainder did not. If a policy clearly states that cookies are not PHI, then the act of attaching a P3P policy to them becomes meaningless, since non-PHI legally does not fall within the scope of a privacy policy. In the case of attaching a P3P policy to a non-PHI cookie, the act only makes sense if the cookie acts as a housing agent and the policy is used in the validation of all forms submitted via the healthcare company's pages, which was not what we observed.

Finally, the use of *machine-readable policies*, e.g. specified using P3P, on healthcare websites and generally in healthcare information systems, is currently unsatisfactory. In fact, none of the organizations in our survey has a P3P version of their notice of privacy practices available on the website. The benefits of better policy design and analysis that a machine-readable policy is envisioned to provide can be manifested in the healthcare sector when the notices of privacy practices are actually posted and utilized in this form on the healthcare organization's website and in their systems. At its core, the use of machine readable policies will force entities to define less ambiguous privacy policies, which would be a huge victory for patients.

Summary

The overall message from our survey was that even though the privacy policies cover enough ground to enable healthcare organizations to arguably claim *regulatory compliance*, they are not adequate to communicate understandable privacy practices to the patient or provide adequate privacy safeguards. We believe that, while using artifacts such as broadly-defined purposes, exception-based accesses and umbrella authorizations may still allow the organizations to claim regulatory compliance, organizations should strive to do better. It is only a matter of time before gaining patient confidence and trust with regards to privacy concerns plays a more significant role.

We alluded to the fact that the levels of privacy protection achieved by a policy depend on its design and enforceability. We observe that the design of a HIPAA-inspired policy hinges primarily on the limited use and disclosure provision, which enables the proactive, fine-grained protection of personal health information. We address how the state of the art can be improved for this provision in the next section.

PRIVACY MANAGEMENT ARCHITECTURE

The goal of PRIMA (**PRI**vacy **M**anagement **A**rchitecture) (Bhatti & Grandison, 2007) is to enable the design of policies with better limited use and disclosure rules. Given the significant investment in healthcare infrastructure and compliance, a key design requirement for privacy-enhancing technology that seeks to bridge the disparity between law and practice is that healthcare information systems must be transformed, with the least possible impact, to an enhanced state of protection.

The task of privacy management in the healthcare industry is complicated by the methods of healthcare services delivery. The clinical workflow in healthcare organizations is fundamentally different from the business process workflow in commercial organizations because of the tremendous amount of human involvement at every step of the way. While it is possible to introduce privacy controls such as auditing and access control between transactions in a business process workflow without compromising its integrity, the same cannot be said about clinical workflows.

For example, a physician routinely takes notes on a piece of paper, and then hands it over to a nurse or other medical staff to input it into the computer system. Whether or not the nurse or medical staff is authorized to view private notes of the doctor becomes a secondary concern in the interest of carrying out the clinical workflow (i.e. delivering care), which would otherwise be impeded if the physician is burdened with the task of typing in the information himself. Similar situations exist when a new patient is admitted to a ward, or is brought to the emergency department, and the *on duty* assistant needs to take (possibly sensitive) notes or retrieve personal information about the patient. Therefore, there is an even stronger requirement for privacy technology that is used in healthcare workflows. This requirement is to ensure that privacy controls are seamlessly embedded into the clinical workflow without impeding the delivery of care.

The problem addressed by PRIMA is how to improve the design of use and disclosure rules in the policy by leveraging the audit results of actual access and disclosure instances, and analyzing those instances that were not explicitly covered by the existing rules in the policy, i.e. accesses allowed through the use of exception mechanism. Improving the policy increases its *coverage*, which is the ratio of the number of accesses and disclosures addressed by the policy to the total number of access and disclosure instances that were recorded in the system, including exceptions. This is particularly important because progress in the specification of privacy policy may not be a pressing national issue, but the need to ensure that the policy embodies what is being done is very pressing for healthcare. Thus, we view this process of policy refinement to increase policy coverage as a critical phase for innovation in health information systems.

There are four main concepts underlying PRIMA (Figure 1). *Stakeholders policies* define the privacy policies of the involved parties. We acknowledge that the patient is the ultimate stakeholder. However, in this context, we expand its scope to include practitioners, payers, etc. *Privacy controls* are the technology components that embody the requirements that need to be enforced in the clinical workflow. *Deployment* is the process that integrates or embeds the privacy controls into the clinical workflow. *Refinement* is the process of continually refining the definition of privacy controls based on audit analysis. In the following section, we discuss each of these concepts in greater detail.

Stake Holder Policies

The multiple stakeholders in a clinical workflow (patient, physician, medical staff, insurance, pharmacy, etc.) may have their own rules and preferences regarding the use and disclosure of protected patient information. These policies need

to be studied in order to be able to specify the privacy policy that applies to the clinical workflow. This can be viewed as a bottom-up approach to policy specification. We expect that a comprehension of the domain and environment will lead to more informed initial policy.

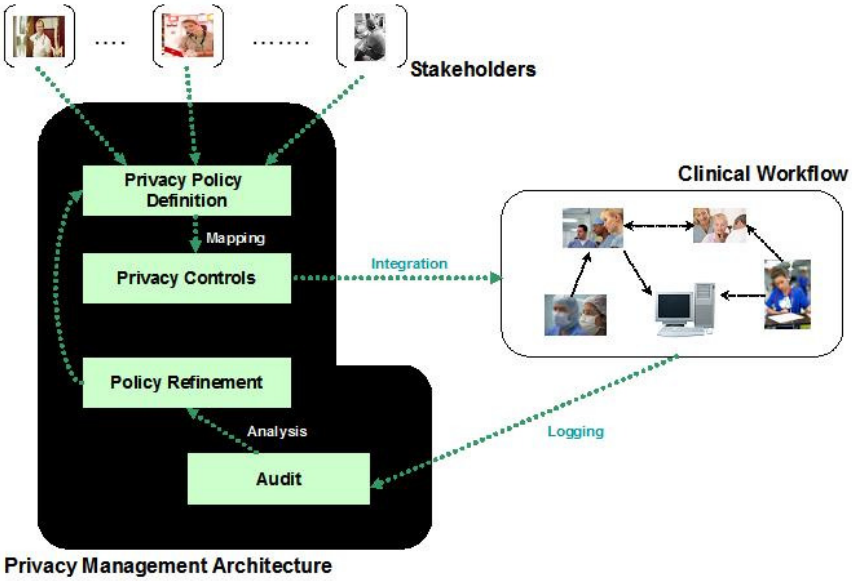

Figure 1: System Architecture

Privacy Controls

The privacy controls of the system in question are based on the privacy policy designed in the previous step. This PRIMA module determines the components, which are to be integrated into the workflow, that enforce the policy specifications. For example, if Juanita disallows the disclosure of her medical record to a medical staff *B* for purpose *Y*, such a rule demands that information retrieved by *B* from the medical database needs to be controlled.

Deployment

The goal of the PRIMA deployment module is to provide a staged deployment of privacy controls by integrating them within the clinical workflow without impeding normal clinical activity. Our survey uncovered that current practice typically allows data access and disclosure against broadly-defined purposes and umbrella authorizations as defined in the policy; and enables the subversion of disclosure controls so that care can be delivered, i.e. allows exception-based access control. Here is a stereotypical example. When *Chiaki*, who is a medical practitioner at Jenkins Community Hospital, requests a record *R* for a patient *P* for purpose *X* that *Chiaki* does not have explicit authorization to view, the system normally provides the information and creates an entry in the audit log that exception access was given. Today, these logs are only used when someone raises a red flag about the improper disclosure of their data. We propose that these logs be used more proactively to close the coverage gap.

Refinement

We purport that if the generated logs are carefully studied, then one can identify instances where certain purposes correlate strongly with a certain category of users wishing to access certain kinds of data. The idea is that those purpose and user category combinations should gradually be incorporated into the policy design so that future such instances should no longer be marked as exceptions. Additionally, the identified purpose definitions and categories of users should be advertised and notified to the patient. Overall, the system moves toward the goal of improving policy coverage.

The Framework for Refinement The audit logs of a covered entity may contain different kinds of information. There may be data on attempts to break into the system, i.e. possible violations or data breaches, or information that represents undocumented, informal clinical practice. Thus, the process of distilling useful facts from the logs requires multiple steps:

1. Prune - This is an optional phase, depending on the extract algorithm used to find informal clinical practice patterns in the log. The overall goal is to reduce the number of artifacts that must be examined in the extract phase

by separating, as much as possible, useful exceptions from violations. This is an area that requires further examination from the research community.

2. Extract - In this phase, an algorithm is applied to the data source to extract patterns that could possibly be incorporated into the policy. There are two types of algorithms that could be used here. The first is a simple matching algorithm that looks for the number of occurrences of term combinations and returns those with high frequencies. This approach assumes that the data source contains only useful exceptions, and thus a pruning algorithm was performed beforehand. The second type of algorithm is a richer data mining algorithm that not only examines the text, but also incorporates information about the relationships between the artifacts being examined. This class of algorithms does not make any assumptions on the contents of the data source and is assumed to contain functionality that reduces the probability of violations appearing in the returned result set. This is another fertile research area. Both classes of algorithms should return the patterns with associated ratings, which indicate the level of usefulness. This measure could be a function of the frequency of the pattern amongst other things. It should also be noted that extraction algorithms must be tailored to the environment in which they are deployed.
3. Filter - Not all the patterns produced from the extraction phase will be appropriate for inclusion into policy; and eventually into the clinical workflow. In fact, some patterns may represent behavior that needs to be stopped. In this phase, either a human is asked which patterns are worthy of inclusion or a program is designed that automatically includes patterns into the policy based on threshold values for the usefulness measure. Even though refinement is an ongoing process, we assume that there has to be a training period, where a reasonable amount of information is collected in the audit log. This training period is totally dependent on the particular healthcare entity deploying the system.

Bhatti & Grandison (2007) provide a use case scenario that demonstrates how PRIMA would enable enhanced privacy protection. The ultimate value of PRIMA lies in its ability to utilize information in the way medical practitioners use data to inform improvements in the policy design, and in the fact that it helps to reduce the number of exceptions over time; since they can become policy rules. As a result, the limited use and disclosure rules, as refined using PRIMA, enable improved privacy protection for the patient. Other technology that would empower patient privacy with regards to the limited use and disclosure provision of HIPAA are: 1) support for fine-grained, purpose-based disclosure controls, 2) support for electronic patient designation of representatives, 3) conflict resolution technology to address multiple policies, state and federal laws, organizational directives, etc., and 4) support for sophisticated data retention and recovery policies. Appendix B contains our thoughts on the key technology needed to enable the other HIPAA Privacy Rule provisions.

CONCLUSION

In this paper, we examined policies from 20 healthcare companies and analyzed the trends with respect to the HIPAA Privacy provisions and to patient privacy expectations. Our survey reveals that the healthcare industry is in need of improved practices. Current practice allows companies to claim compliance with regulation, but fell short of what patients expect with regards to their privacy. We then focused on a particular problem that currently exists in healthcare companies and presented a solution, based on policy refinement, that enables better design of policy by leveraging information from the clinical workflow.

ENDNOTES

[1] The Health Insurance Portability and Accountability Act is commonly referred to as HIPAA.
[2] The Health Information Technology for Economic and Clinical Health Act is commonly referred to as HITECH
[3] The American Reinvestment and Recovery Act is commonly referred to as ARRA.
[4] The Personal Information Protection and Electronic Documents Act is commonly referred to as PIPEDA
[5] P3P Preference Exchange Language is commonly referred to as APPEL

REFERENCES

- Privacy Rights Clearinghouse (2009). A chronology of data breaches. Retrieved November 27, 2009 from http://www.privacyrights.org/ar/ChronDataBreaches.htm.
- U.S. Department of Health and Human Services (1996). Health Insurance Portability and Accountability (HIPAA) Act. Retrieved November 27, 2009 from http://www.hhs.gov/ocr/privacy/hipaa/administrative/privacyrule/adminsimpregtext.pdf.

- U.S. (2009), Health Information Technology for Economic and Clinical Health (HITECH) Act. Retrieved November 27, 2009 from http://waysandmeans.house.gov/media/pdf/110/hit2.pdf
- U.S.A. (2009), American Recovery and Reinvestment Act (ARRA). Retrieved November 27, 2009 from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_bills&docid=f:h1enr.pdf
- Robert Pear (in press). Warnings over privacy of us health network. New York Times, February 18, 2007.
- L. Rostad and O. Edsburg (2006). A study of access control requirements for healthcare systems based on audit trails from access logs, In the proceedings of the Annual Computer Security Applications Conference (pp. 175-186). Miami Beach, Florida, USA: IEEE Computer Society.
- Rebecca Wong (2009). An overview of data protection laws around the world. Retrieved November 27, 2009 from http://pages.britishlibrary.net/rwong/dpa.html.
- Office of the Privacy Commissioner of Canada (2000). Personal Information Protection and Electronic Documents Act (PIPEDA). Retrieved November 27, 2009 from http://www.priv.gc.ca/legislation/02_06_01_e.cfm.
- Japanese Ministry of Internal Affairs, Communications Information, and Communications Policy (2003). Personal data protection law. Retrieved November 27, 2009 from http://www.kantei.go.jp/jp/it/privacy/houseika/hourituan/index.html.
- HIMSS (2009), Privacy and Security Toolkit. Retrieved November 27, 2009 from http://www.himss.org/CPRIToolkit/html/4.11.html.
- U.S. President's Information Technology Advisory Committee (PITAC) (2004). Revolutionizing Health Care Through Information Technology. Retrieved November 27, 2009 from http://www.nitrd.gov/pitac/reports/index.html.
- L. Cranor, M. Langheinrich, M. Marchiori, M. Presler-Marshall, and J. Reagle (2002). The Platform for Privacy Preferences 1.0 specifications. W3C Recommendation. Retrieved November 27, 2009 from http://www.w3.org/TR/P3P/.
- S. Byers, L. F. Cranor, and D. Kormann (2003). Automated Analysis of P3P-enabled web sites, In the proceedings of the International Conference on Electronic Commerce (pp. 326-338). Pittsburgh, Philadelphia, USA. ACM Press.
- G. Karjoth, M. Schunter, and M. Waidner (2002). Platform for Enterprise Privacy Practices: Privacy-enabled management of customer data, In the proceedings of the International Workshop on Privacy Enhancing Technologies. San Francisco, California, USA. Springer Verlag.
- M. Schunter, E. V. Herreweghen, and M. Waidner (2002). Expressive Privacy Promises how to improve the Platform for Privacy Preferences (P3P), In the proceedings of the W3C Workshop on Future of P3P. Dulles, Virginia, USA. Retrieved November 27, 2009 from http://www.w3.org/2002/p3p-ws/pp/ibm-zuerich.pdf
- N. Li, T. Yu, and A. I. Anton (2003). A Semantics-based Approach to Privacy Languages. Technical Report. Center of Education and Research in Information Assurance and Security. Retrieved November 27, 2009 from http://www4.ncsu.edu/~tyu/pubs/p3p-csse06.pdf
- McDonald, A., Reeder, R., Kelley, P. and L. Cranor (2009). Comparative Study of Online Privacy Policies and Formats. In the proceedings of the Annual Privacy Enhancing Technology Symposium (PETS). Seattle, Washington, USA. Springer Lecture Notes in Computer Science series (Volume 5672).
- Jensen, C., Potts, C. and C. Jensen (2005). Privacy practices of Internet users: Self reports versus observed behaviour. *International Journal of Human-Computer Studies*, 63, 203-227.
- I. Anton, J. B. Eart, M. W. Vail, N. Jain, C. M. Gheen, and J. M. Frink (2007). HIPAA's effect on web site privacy policies. IEEE Security and Privacy, 5(1), 45–52.
- Ross Anderson (1996). A security policy model for clinical information systems, In the proceedings of the IEEE Symposium on Security and Privacy. Oakland, California, USA. IEEE Press.
- IHE (2006). The Patient Care Coordination Technical Framework: Basic Patient Privacy Consents, Supplement 2005-2006. Retrieved November 27, 2009 from http://www.ihe.net/Technical_Framework/upload/IHE_PCC_TF_BPPC_Basic_Patient_Privacy_Consents_20060810.pdf.
- Bernd Blobel (2004). Authorisation and Access Control for Electronic Health Record systems. *International Journal of Medical Informatics*, 73(3), 251-257.
- M. Hochhauser (2001). Lost in the fine print: Readability of financial privacy notices. Retrieved November 27, 2009 from http://www.privacyrights.org/ar/GLB-Reading.htm.
- Graber, M. A., D'Alessandro, D. M. and J. Johnson-West (2002). Reading level of privacy policies on internet health web sites. *Journal of Family Practice*.
- M. J. Ball, C. Smith, and R. S. Bakalar (2007). Personal Health Records: Empowering Consumers. *Journal of Healthcare Information Management*, 21(1).
- Pollach (2007). What's wrong with online privacy policies? Communications of the ACM 30(2), 103-108.
- Rafae Bhatti, Tyrone Grandison (2007). Towards Improved Privacy Policy Coverage in Healthcare Using Policy Refinement. In the proceedings of the 4th VLDB Workshop on Secure Data Management. Vienna, Austria. Springer.

APPENDIX A - HIPAA PROVISIONS AND ALLOWED USES/DISCLOSURES

Provisions

The five key principles of the HIPAA Privacy Rule are:

1. **Notification** - Each covered entity must provide all patients with a notice of its privacy practices, describing the ways in which the covered entity may use and disclose protected health information. It must notify patients of their rights under the law, the covered entity's duties regarding PHI, and their right to complain to the Department of Health and Human Services (HHS) and the covered entity.
2. **Authorization and Consent** - Covered Entities must obtain written authorization from the patient for disclosures that are not for treatment, payment, or health care operations or otherwise permitted by the Privacy Rule. For instance, a covered entity must obtain written patient authorization prior to disclosing any personal health information for marketing purposes subject to limited exceptions.
3. **Limited Use and Disclosure** - Covered Entities must make reasonable efforts to use, disclose, and request only the minimum amount of protected health information necessary to accomplish the intended purpose (known as the *minimum necessary* principle). They also must develop and implement policies and procedures that restrict access and uses of protected health information based upon the specific roles of the members of their workforce.
4. **Auditing and Accounting** - Patients have the right to an accounting of all disclosures of their protected health information by a covered entity the covered entity's business associates for a maximum of six years past. The covered entity does not have to account for disclosures: (a) for treatment, payment, or health care operations; (b) to the patient or its representative; (c) to individual's involved in patient's health care or payment; (d) pursuant to authorization; (e) of a limited data set; (f) for national security or intelligence purposes; (g) for inmates in lawful custody; (h) incident to otherwise permitted uses or disclosures.
5. **Access** - Patients have the right, under most circumstances, to access the covered entity's designated record set, which contains medical, billing, and other information used to make decisions about that individual patient upon patient request. Covered entities must amend information that is inaccurate or incomplete.

Allowed Uses and Disclosures

Under the HIPAA Privacy Rule, PHI can be used or disclosed:

1. to the individual,
2. for treatment, payment, or health care operations,
3. when patients are given the opportunity to agree or object, i.e. in the creation of patient directories, etc.,
4. when it is incident to an otherwise permitted use or disclosures,
5. when it is in the public interest and benefit activities, e.g. law enforcement, public health activities, domestic violence, health oversight activities, judicial/administrative proceedings, law enforcement, threat to health or safety, IRB-approved research, and
6. when a *limited data set* is used for research, public health, or health care operations.

Generally, there are no restrictions on the use or disclosure of de-identified information, which is defined as information with no identifiers. Under HIPAA, an acceptable level of de-identification has been reached under the following conditions:

1. *Limited Data Set* - The covered entity has removed a list of sixteen ( ) direct identifiers (named in the Privacy Rule) and a data use agreement is in place.
2. *Safe Harbour* - Removal of base 16 identifiers plus (1) all geographic identifiers except first three digits of zip code (if more than 20,000 people); (2) all elements of dates except year; (3) any other unique identifying number, characteristic, or code, except as permitted for re-identification purposes provided certain conditions are met.
3. *Alternate De-Identification Method* - A person with appropriate knowledge of generally accepted statistical and scientific methods determines there is a very small risk that the information could be used, alone or in combination with other available information, to identify the data subject; and the covered entity documents the methods and results of this determination.

APPENDIX B - RECOMMENDATIONS

Technology developed to address the recommendations below should be designed with two imperatives in mind: 1) solutions must be efficient, in terms of system resources required and execution timescales, and 2) solutions must be transparent, i.e. allow for seamless and non-intrusive inclusion into existing healthcare infrastructure and workflows.

1. **Notification**: Emerging healthcare solutions should address how one tracks and triggers (technology) events that lead to notification and how one ensures receipt and includes the receipt of notification into the electronic system.
2. **Authorization and Consent**: Current posted notices of privacy practices and policies do not facilitate automated analysis. The first step in enabling this is the creation of systems and mechanisms for the electronic representation of patient authorizations and consents. Going a step further, one needs to identify the appropriate granularity for authorizations and consents and develop tools that collect and proactively use these authorizations and consents in healthcare information systems.
3. **Auditing and Accountability**: Auditing technology for all healthcare activity (i.e. disclosures, modifications, etc.) is sorely needed. The audit results must be presented in order of relevance and summarized to be meaningful to the person issuing the audit. Technology for ensuring tamper-resistance of audit logs and the development of new audit structures and mechanisms that have small storage footprints and performance impact is critical. Enhanced functionality on top of audit logs should be defined to ensure better data quality, analyze behavior and provide other valuable insight.
4. **Access**: Patients need technology that enables the secure access and download of their electronic health record.

These are a few of higher level technology contributions that can be made to foster better healthcare privacy practices.