Regulatory Privacy Protection for Biomedical Cloud Computing

Y. Tony Yang¹, Kari Borg²

¹Department of Health Administration and Policy, George Mason University, Fairfax, USA; ²Department of Ophthalmology, Polyclinic, Seattle, USA.
Email: ytyang@gmu.edu

Received August 7th, 2012; revised September 4th, 2012; accepted September 15th, 2012

ABSTRACT

This article provides background information on biomedical cloud computing. It examines the privacy concerns that arise from the use of biomedical cloud computing services and then surveys the current state of regulatory privacy safeguards for patients and consumers of these services both in the US and abroad. Finally, it identifies opportunities for legal and technological mechanisms to be implemented or reinforced so that patients and consumers are not forced to lose control of their information when they use biomedical cloud computing services.

Keywords: Privacy; Law; Regulation; Biomedicine; Cloud Computing

1. Introduction

Advances in technology and progressions in web-based software have led to the ability to share information with ease and efficiency. In relation to health care, this has created the ability to access patient’s electronic medical records from virtually anywhere. Health care professionals now have the capability to save more lives with faster access to pertinent medical information. Medical researchers have also been utilizing this technology to compile data as well as allow other users the right to use the data for further exploration. These researchers are now gathering data from a virtual storage facility where they are able to store information and permit other researchers to access it as well. This form of storing data is known as “cloud computing”: operates through a third party organization that monitors the use, security, and accessibility of the stored data. While this has created major innovations in health care research, it brings up a serious concern regarding patient privacy and confidentiality. Records of patients’ personal medical histories, as well as other identifying data, have a high risk of being abused when stored in the cloud. Patient data is now in a complex, virtual world that is constantly threatened by hackers and internal breaches in security. This ongoing issue has created debate on how to properly protect patient privacy while still allowing professionals to have the access they need to important data.

2. Biomedical Cloud Computing

Biomedical data is increasingly being stored on the “cloud”. The cloud allows for multiple users to access a shared data for research much faster than ever before. Data that used to take researchers months to gather can now be extracted from the cloud in minutes or even seconds (Crawford [1]). One example of a biomedical cloud is that of Ohio State University. They have created a cloud system known as the Translational Research Informatics and Data management grid or “TRIAD” (Crawford [1]). This software was developed to allow most forms of biomedical data to be uploaded and stored in this cloud, with software then interpreting the data into a standard universal language (Crawford [1]). “When it comes to biomedical research, you have the digital equivalent of the Tower of Babel. One piece is written in French. And another is written in Russian. And maybe a third component is in Chinese,” says Philip R. O. Payne, a researcher and chair on the department of biomedical informatics at the Ohio State University (Crawford [1]). Researchers can now extract, upload and share data without the need to spend time and money on translating the data into their own “language”.

United Health Inc. is also creating a new cloud aimed at merging the financial aspect of healthcare. For example, patients will have their billing combined from various locations regarding the same diagnosis or procedure that was done (Mathews [2]). UnitedHealth is also aiming at creating a cloud with data pooled from multiple electronic health records to allow doctors to see if their patient was seen elsewhere, as well as for emergency rooms seeking the entirety of their patients’ charts while treating them (Mathews [2]). These private health record
systems are believed to improve doctor-patient communication, make understanding and use easier, and reduce the risk of medical errors (Carrión, Alemán & Toval [3]).

Electronic Health Records (EHRs) that are accessed by multiple organizations or even different locations, rely on some form of off-site, or cloud, computing (McCarthy [4]). Although the benefits of these systems to the medical research world are significant, the importance of a solid framework to manage the security of the data cannot be understated. While all of these interfaces have privacy policies in place, the virtual nature of the data means that it is not primarily protected by physical devices such as locks or buildings. The cloud’s virtual storage lockers are at risk for being hacked into regardless of the privacy policies put in place to regulate who can access personal identifying information.

3. Privacy Risks in the Clouds

Organizations that manage data clouds are required to have privacy policies in place to create a secure network for their users and for the patients whose data has been extracted. One of the top concerns with storing sensitive medical data is risking the patients’ identifiable information (PII). Medical researchers keep data concerning patients’ medical diagnosis, family history, HIV status, etc. It also includes patients’ social security numbers, addresses, date of births, among many other records. One form of preventing identifiers from being stolen and abused is to have the system erase all unnecessary personal data when information is being pulled and correlated with other figures. For example, TRIAD has adopted an interface that allows researchers to connect tissue samples with medical records that will de-identify the record making the correlation completely anonymous (Crawford [1]). By making it difficult for users to attain identifying information, a significant amount of potential identity theft is deterred. Biomedical cloud providers must follow strict guidelines in regards to patients’ privacy by the Health Information Portability and Accountability Act (HIPAA). For instance, there are eighteen types of identifying information that must be removed before biomedical data can be shared with other parties without the consent of the patient (Klein [5]). Some additional restrictions include being sure the server for the network is within the United States and must have physical means of protection such as cages, back-up power, security guards, etc. They must also follow guidelines set forth by the Privacy Act of 1974 which protects health information and allows patients the right to see and change information on their records (Osterhaus [6]). In 2009 the Health Information Technology for Economic and Clinical Health (HITECH) [7]. Act was signed to encourage meaningful use of health care technology and “strengthen the civil and criminal enforcement of the HIPAA rules”. The HITECH act made the regulations present in HIPAA applicable to not just healthcare organizations, but also the cloud service providers that they utilize, closing an immense loophole in security (Delgado [8]).

A group called the Cloud Security Alliance is a non-profit organization “with a mission to promote the use of best practices for providing security assurance within Cloud Computing, as and to provide education on the uses of Cloud Computing to help secure all other forms of computing”. The Cloud Security Alliance has determined the top seven threats to cloud computing as (Klein [5]): 1) Abuse and nefarious use of cloud computing; 2) Insecure application programming interfaces; 3) Malicious insiders; 4) Shared technology vulnerabilities; 5) Data loss or hijacking; 6) Account, service and traffic hijacking; 7) Unknown risk profile.

Users require that cloud providers be trustworthy for storing their data. To create trust there must be security in place for all of the sensitive information. There are a number of ways in which to address this complex concern. One is by examining three aspects involved in the security of these virtual storage facilities: preventative, detective, and corrective controls (Ko, et al. [9]).

Preventative controls are the measures put in place to prevent breaches from occurring in the first place (Klein [5]). These can include encryption technology, firewalls, lists of approved users, etc. Many of the controls specified by HIPAA fall under this category, and all can be readily implemented by cloud service providers to suit the specialized needs of the health industry (Martin [10]). Each layer of protection that can be added must be evaluated for both its costs and benefits. For instance, one possible method of preventing data loss is having dedicated and distinct servers for each client. While this increase in isolation bolsters security by minimizing “accidental” thefts when the target was on the same server, this approach increases the costs to the server provider. Additionally, this diminishes the benefit of having the information stored in multiple locations, which minimizes loss because of hardware malfunctions (Palanzi [11]).

One more sophisticated, albeit partial, means of preventing unnecessary transfer of information is through a cloud-based technique called Virtual Machines (VMs). Virtual machines are minimal terminals that load their operating system from a server, either through local infrastructure or the internet (Grossman & White [12]). Most, if not all, processing is performed by the server. Rather than transferring the entire contents of the file system over the internet, only the pixels displayed on the screen are transmitted. In the case of viewing of medical images (which range from 2 - 500 mb) this approach drastically reduces the amount of bandwidth necessary to view and manipulate files remotely. The nature of VM systems prevents much of the unnecessary transferring of sensitive
patient information from the data center to a terminal when accessing EHRs. Additionally, this approach is generally more cost effective, reliable and easier to configure than traditional approaches that rely on the installation and maintenance of one operating system install per computing system (Philbin, Prior & Nagy [13]). Furthermore, they provide a uniform computing environment whether accessed from work, home or on the road.

Detective controls consist of procedures taken to identify risks in the system (Klein [5]). They include guidelines for security administration (such as routinely conducting risk analyses and implementing policies and procedures to address vulnerabilities), screening and educating the workforce, and activity audits (Schweitzer [14]). Corrective controls are actions to resolve breaches that have already occurred and need to be fixed (Klein [5]). In having the cloud provider monitor these controls, trust can flourish by the users knowing which and how security measures are being applied to secure their information. Creating a strong structure for managing these controls is the key for adherence to the privacy policies governing health data clouds.

Predictive, detective, and corrective controls are applicable to both internal and external threats of information privacy. Many, if not most, breaches of sensitive data are perpetrated by internal threats, i.e. employees working at either the healthcare organization or cloud service provider. There are two important themes underlying internal security threats: accountability and auditability.

Of the seven threats outlined by Ko, et al. [9], all but the 4th and 6th risk can be addressed through increasing accountability and auditability of the cloud providers (Klein [5]). They state that to create trust in the cloud provider, the provider must address these issues to show its users the measures they are taking to secure data and private information (Klein [5]). Consumers who are storing data in the cloud must have full trust in the cloud provider due to the fact that they are keeping substantial medical and personal identifying information in this system (Carrión, Alemán & Toval [3]). This creates incentive for the provider to show its users how they are addressing these privacy issues. Cloud providers must be able to provide accountability for any problems that arise within their system and display this to their users.

They must also have their system logs and design clearly presented for their customers, so that users can feel secure in knowing the activity and protection of their accounts—Something that many cloud services currently fail to do. There are no means to audit a cloud structure if there is no monitoring system for users so that they can survey how and when their data is stored and accessed. Accountability and auditability are crucial for health organizations to feel comfortable in fully utilizing the potential powers of the cloud; thorough logging is the cornerstone necessary for both (Pearson [15]).

Accountability in a cloud system is necessary for determining which party is responsible when a security breach occurs. If information is taken by an employee of the health institution, liability typically falls similarly. However, if the breach is perpetrated by an employee of the cloud service provider or an external agent, responsibility should lie with the service provider. A retrospective analysis of system logs can typically reveal the origin of a breach, but how to efficiently (in terms of size) and effectively (in terms of preventing future breaches) log health information is still an open question. Many clients that use VM are not fully aware of the linkages between virtual and physical servers, relationships between virtual and physical server locations, and how files are written into both virtual and physical memory addresses (Ko, et al. [9]). The complexity inherent in virtual machine set-ups can be a difficult challenge in terms of determining the origin of a breach. Without proper measures implemented ahead of time, assessing liability may be impossible—Put simply, a legal nightmare.

Auditability refers to the relative ease of auditing a system or an environment. A system lacking in auditability has poorly maintained records (Soma, et al. [16]). The complex nature of cloud data exchange necessitates a multilayer approach for modeling information exchanges. At the system level, there is the operating system, the file system, and the network, as well as the logs corresponding to the read/write operations between them. At the data layer, there is the logging of each detail regarding a single file: its origin, the location(s) where it is stored, which accounts what permissions to access to it, and if/when it will expire. The most important bit of information in logs is the time-stamped record of account accessing and interacting with files. So long as these details are both comprehensive and properly protected with encryption, establishing liability is usually possible. Finally, assessment of the workflow level reveals the robustness or weakness of the control regarding user/terminal interaction: are employees only accessing the minimal amount of information they need to accomplish a task? Through workflow analysis, problem areas and security loopholes in the cloud are removed or rectified and control and governance of the cloud processes are improved.

Kirchberg et al. have suggested that a file-centric perspective should take prevalence in health-related cloud computing security (Ko, et al. [9]). Network logs that trace the life cycle of files (i.e. creation, modification, duplication and destruction) within clouds best enable accountability and auditability. Additionally, event data regarding network activities and actor data (the person or computer component that triggered an event) are necessary for proper representation of the file life cycle. Estab-
lishing different network zones is perhaps the best way to organize network logs. A cloud service provider, for instance, can designate its own network as internal and the network addresses of the health organization as a safe zone. Data that is relayed to addresses besides these zones can be flagged for review by an automated or human supervisor. In the future, file types could be designed to include some space for the local storage of pertinent meta-information (such as read/write histories), making the logging needs of the distributed cloud system less complex (Ko, Lee & Pearson [17]).

The lack of transparency within the cloud service provider’s system is the greatest concern for prospective cloud users. One survey found 88% of potential cloud customers were concerned about who would have access to their data (Fujitsu Research Institute [18]). Some EHR cloud providers, such as Microsoft Health Vault, fail to provide thorough accounts of data storage and access for their users. All contractual relationships between cloud service providers and medical organizations should be considered with care; especially in cases when all of the relevant information for assessing liability may not be accessible by both parties. For this, and other security reasons, the organization needs to negotiate a strong contract with the cloud service provider featuring compliance with HIPAA, security controls, and auditability (Witt [19]). If these requirements can be met, the benefit from shared data sets between health researchers will be immense. In the near future, biomedical data can be mined for statistical patterns between an individual’s genome, therapy, and conditions (Grossman [12]). Increasingly open and shared data means greater statistical power in detecting predispositions to diseases early.

4. Regulatory Protection for Privacy

4.1. Privacy Law in the US and Abroad

Biomedical enterprises operating in the United States need to consider HIPAA (45 C.F.R. §§ 160 - 164) and the Electronic Communications Privacy Act (ECPA) (18 U.S.C. §§ 2701 - 2712), a federal statutory law that supports specific protections for electronic communications (in transit or in storage) to supplement any protections offered by the Fourth Amendment of the United States Constitution. Such laws require privacy and data security obligations.

European-based biomedical enterprises, as well as entities working with providers in or with infrastructure in Europe, however, need to take into account the expansive requirements under local omnibus data protection laws that safeguard all personal information, even basic details like business contact information. These requirements can involve informing employees, customers, or other individuals about the outsourcing and processing of their data; obligations to consult with works councils before outsourcing employee data; and registering with local data protection authorities (Lyon [20]).

4.2. Requirements for Data Security

Even if an enterprise is not subject to these types of privacy laws, it will want to ensure safeguards for personal information covered by data security and breach notification laws. In the United States, these laws focus on personal information such as social security numbers, driver’s license numbers, and credit or debit card or financial account numbers. One of the key safeguards is encryption because many (although not all) of the US state breach notification laws provide an exception for encrypted data. In contrast, many other countries require protection of all personal information, and do not necessarily provide an exception for encrypted data. Consequently, companies operating outside of the United States may have wider-reaching obligations to protect all personal information. While data protection obligations vary considerably from law to law, both US and international privacy laws commonly require the following types of safeguards: 1) Conducting appropriate due diligence on providers; 2) Restricting access, use, and disclosure of personal information; 3) Establishing technical, organizational, and administrative safeguards; 4) Executing legally sufficient contracts with providers; and 5) Notifying affected individuals (and potentially regulators) of a security breach compromising personal information (Delgado [8]).

The topic of data security in the cloud has received significant attention lately. The National Institute of Standards and Technology (NIST) has finalized its first set of guidelines for managing security and privacy issues in cloud computing in early 2012. Guidelines on Security and Privacy in Public Cloud Computing (NIST Special Publication 800 - 144) provides an overview of the security and privacy challenges facing public cloud computing and presents recommendations that organizations should consider when outsourcing data, applications and infrastructure to a public cloud environment (Lyon [20]). The document provides insights on threats, technology risks and safeguards related to public cloud environments to help organizations make informed decisions about this use of this technology. SP 800-144 is geared toward system managers, executives and information officers making decisions about cloud computing initiatives; security professional responsible for IT security; IT program managers concerned with security and privacy measures for cloud computing; system and net-work administrators; and users of public cloud computing services (Delgado [8]).
4.3. Restrictions on Data Transfers Internationally

A growing number of countries, especially in Europe, restrict the transfer or sharing of personal information beyond their borders. These restrictions can present significant challenges for multinational biomedical enterprises seeking to move their data to the cloud. Recognizing these challenges, some providers are starting to offer geographic-specific clouds, in which the data are maintained within a given country or jurisdiction. Some US providers have also certified to the US-European Union (EU) Safe Harbor program, in order to accommodate EU-based customers and patients. However, as the Safe Harbor only permits transfers from the EU to the United States, it is not a global solution. Accordingly, a biomedical enterprise should assess carefully whether the options offered by a provider are sufficient to meet the enterprise’s own legal obligations in the countries where it operates (Delgado [8]).

5. Critical Recommendations for Reinforcement of Privacy Protection

5.1. Legal Reform

The law has long recognized the importance of privacy. Existing statutory privacy law, however, needs a technological upgrade. For example, ECPA should, but does not clearly define the statutory protections applicable to cloud computing services. ECPA does not distinctly express whether documents stored with many cloud computing services are protected at all. ECPA, as currently written, provides protection where content is stored with a service “solely for the purpose of providing storage or computer processing services to such subscriber or customer, if the provider is not authorized to access the contents of any such communications for purposes of providing any services other than storage or computer processing.” It is not clear whether sites that provide collaboration and sharing functions or employ a targeted advertising business-model based on information contained in documents are covered by this clause. Even if ECPA does cover cloud computing records in a specific situation, the protections that it provides are insufficient to properly safeguard the privacy of sensitive documents being stored with cloud computing services. Beyond ECPA, there are questions about whether other specific privacy laws or regulations fully protect consumers of cloud computing services. For instance, the HIPAA is designed to protect the privacy of health records. HIPAA applies to health care providers, health care clearing-houses, and health plans (insurers). But, it is not clear whether HIPAA protections apply to cloud computing services that store consumer and patient health records. Similarly, does the Video Privacy Protection Act (18 U.S.C. § 2710), which provides statutory protection for video rental records and “other similar material,” protect records of audiovisual material shared or retrieved through a cloud computing service?

Therefore, as biomedical cloud computing continues to evolve, it is imperative that privacy laws and policies are updated so that consumers and patients have the clarity needed to make informed choices and feel confident that their personal information is being protected (Weissberger [21]). To this end, Congress might consider implementing some of the enforcement provisions of the False Claims Act to strengthen HIPAA. This would allow a private right of action measure similar to the FCA’s qui tam actions (Palanzi [11]). Furthermore, by increasing and modifying the fine structure for services that store large numbers of electronic medical records on cloud networks, the compliance of cloud companies to meet new regulatory standards may be financially incentivized (Palanzi [11]).

5.2. Contractual Protections for Cloud-Bases Services

Cloud providers must be trusted to maintain the integrity and security of the data they are tasked with storing. Establishing this trust can be achieved by means of a contract which allows the customer the option of strong legal recourse in the event of a data breach. The contract will govern the conditions upon which data is given to the cloud computing provider and should be drafted in consultation with a local attorney or other third party specializing in data security laws (Harshbarger [22]).

Contractual provisions for cloud based services should include an insurance policy for damages resulting from compromised data, as well as a stipulation of service levels. The former demonstrates that the cloud provider has the finical ability to pay a breach of data while the latter clearly delineates the responsibilities on the part of the cloud service provider. Examples of service level stipulations include having the cloud provider state that it will have a ninety-nine percent up-time and have its system free from bugs or other defects (Harshbarger [22]). In addition, the cloud provider could negotiate additional operational specifications which might be unique to the nature of the data the customer needs to be stored. This would give the customer a defined set of expectations in relation to the cloud provider’s handling of sensitive data.

A contract would also allow for increased transparency on the part of the cloud provider. A provision for transparency would oblige the cloud provider to make its data security regime available to the customer so that specific precautions and safeguards are known (Harshbarger [22]). This transparency will increase trust and
5.3. Vigorous Privacy Practices from Biomedical Enterprises

Biomedical enterprises have the opportunity to ardently address much of this patient and consumer concern by following the core principles of the Fair Information Practices (Federal Trade Commission [23]). This means providing meaningful notice about how information is used and to whom it is disclosed, collecting and retaining only the information that is needed to provide services, giving patients and consumers real choice about how any personal information collected about them will be used, properly safeguarding patient and consumer information from disclosure and misuse, and enabling them to control, modify, and delete their own records and accounts (Turock, et al. [24]). Providing patients and consumers with meaningful control and protection for their personal information will help give them the confidence to utilize cloud computing and may also help biomedical enterprises avoid negative press, government investigations, and costly lawsuits (Ozer [25]).

Patients and consumers expect that data stored with a cloud service provider will stay private; providers have a commercial incentive to make sure that it does. By designing a service with technical measures to protect consumers and patients—Tools that allow them to manage and protect their own information, encryption and anonymity protocols to protect information by default, and access controls and data security measures to prevent breaches and inappropriate disclosures—Biomedical cloud computing providers can establish a platform where patients and consumers are in a position to control their own information and can feel more confident banking private content. The key step in giving patients and consumers control is to build a vigorous and operative interface to allow them to be in charge of their own content and records. Consumers and patients should be able to view and control their entire record. Constructing such an interface is much easier if it is part of the design process of the service and not added on as an after thought or in response to consumer and patient demands for greater control and transparency (Pearson & Charlesworth [26]).

Anonymization and encryption can also protect consumers by reducing the risk of disclosure of information that is captured and stored by the service. Anonymization procedures, however, should ensure that data is irrevocably de-identified (Narayanan & Shmatikov [27]). Moreover, creating a solid data security plan protects not only patients and customers but also providers. Data breaches can be disastrous, leading to lawsuits, fines, and lost trust (Soma, et al. [16]). In order to avoid these outcomes, providers should use access controls to prevent unauthorized access to content by both employees and third parties and take additional steps such as promptly deleting data that is no longer necessary in order to reduce the risk of breach. Such practices will help safeguard both patient and customer privacy and the provider’s bottom line. Providing technical measures that protect and secure patient and consumer information may carry both practical and legal significance. Practically, the measures suggested above, and others that may emerge, reduce the likelihood of breach or unnecessary disclosure. In addition, these mechanisms may strengthen the legal positions of both consumers and providers by making it clear that the patient or consumer, and not the provider, is the party with access to and control over any stored content. The more “locks” a provider puts in the patient or consumer’s control, the less likely it is that third parties will be asking providers for the keys (Wayner [28]).

6. Conclusion

As biomedical cloud computing continues to develop and expand, it is critical to reinforce regulatory mechanisms to protect the privacy of consumers and patients. Courts and policymakers need to recognize the realities of modern information storage and satisfy the continued expectations of privacy, regardless of whether the information is stored online or offline. Biomedical enterprises should invest in privacy-friendly technologies and practices that put consumers and patients in control of their own private information. They should also support regulatory reform to update any outdated statutory understandings of online privacy.

REFERENCES

[5] C. Klein, “Cloudy Confidentiality: Clinical and Legal Im-
Regulatory Privacy Protection for Biomedical Cloud Computing


Copyright © 2012 SciRes.