academicJournals

Vol. 10(5), pp. 184-195, 15 March, 2015
DOI: 10.5897/SRE11.862
Article Number:CEE470D51274
ISSN 1992-2248
Copyright©2015
Author(s) retain the copyright of this article
http://www.academicjournals.org/SRE

Scientific Research and Essays

Full Length Research Paper

Sharing healthcare information based on privacy preservation

Asmaa Hatem Rashid* and Norizan Binti Mohd Yasin

Department of Information Science, Faculty of Computer Science and IT, University of Malaya, Kuala Lampur, 50603 KL, Malaysia.

Received 7 May, 2011; Accepted 19 November, 2014

The evolution and development of information technology have facilitated greater sharing of data and knowledge management for the collection of electronic information by data owners such as governments, corporations, and individuals. Therefore, they have created huge opportunities for knowledge management and information retrieval. Recent develoments have helped improve decision making especially in the fields of medical information, research, and public health organization, among others. Recently, the control and sharing of data or knowledge management has received notable attention in research communities. Many approaches have been proposed for different data publishing needs in different fields. The sharing of data needs control and management to ensure system integration. Integration is required especially in the management of patient data to secure sensitive information such as patient identification. Several studies have focused on the management of data in medical applications to ensure system integration. However, the management and sharing of data in different fields may result in misuse of information. Therefore, there is a need to build models or design certain algorithms to manage shared data efficiently and to avoid misuse. The goal is to ensure authenticity of the data system. In the present study, we systematically summarize and evaluate different approaches to control the sharing of data and knowledge management in order to ensure system integration. Moreover, we study the challenges in controlling the sharing of data and clarify the differences and conditions that distinguish the control of sharing of data from other related problems. Finally, we correspondingly propose future research directions in the conclusion.

Key words: Knowledge management, electronic information, information retrieval, decision making integration, authenticity.

INTRODUCTION

The use of information and communication technology (ICT) in healthcare is increasing (Ernstmann et al., 2009) because of its potential to improve the effectiveness and efficiency of healthcare (Kohn et al., 1999). Health information systems (HISs) help ensure that patients immediately receive appropriate treatment. Aggelidis and Chatzoglou (2009) mentioned that the use of information systems in the healthcare sector is widely accepted,

*Corresponding author. E-mail: asmaarashid@siswa.um.edu.my

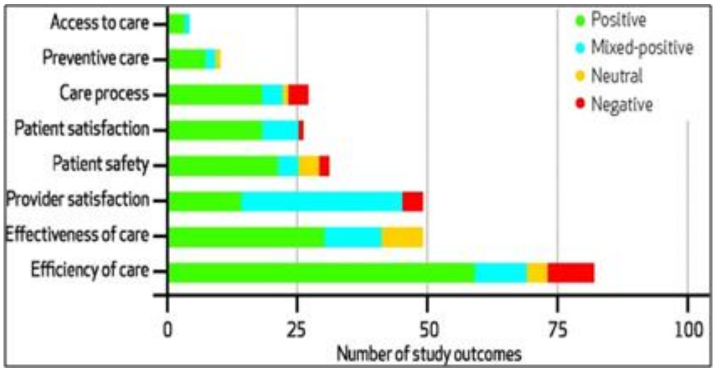

Figure 1. Evaluations of outcome measures of health information technology, by type and rating (Buntin et al., 2011).

particularly in hospitals (Aggelidis and Chatzoglou, 2009). Information systems (ISs) improve the quality of services being provided (Scott, 2007). Researchers reported that the failure of hospitals to adopt new ISs increases inconvenience and loss of the trust among patients (Ammenwerth et al., 2003; Lu et al., 2005). Thus, HISs have gradually replaced traditional hospital procedures (Ammenwerth et al., 2003; Lu et al., 2005), and studies have proposed various frameworks for building trustworthy IS solutions for hospitals.

Healthcare information systems (HISs) in healthcare organizations such as hospitals is important for providing and sharing healthcare information among medical staff, especially physicians and researchers (Yang et al., 2010). In addition, collaboration is an important requirement for HISs (Ahmed and Yasin, 2012). The term “collaboration” in the field of healthcare is defined as the communication that occurs among healthcare practitioners when sharing information and skills regarding patient care (Gaboury et al., 2009; Scandurra et al., 2008; Weir et al., 2011). Furthermore, healthcare information is valuable to many organizations for scientific research or analysis (Chen et al., 2012). Sharing these healthcare data among different organizations can significantly benefit both medical treatment and scientific research in relevant sectors (Hillestad et al., 2005; Wang et al., 2003; Yang et al., 2010). Nevertheless, healthcare data typically contains considerable private information. Sharing this data directly would pose a threat to patient privacy. Thus, developing practical models to balance healthcare data sharing utility and privacy preservation is necessary in order to improve collaboration among physicians (Chen et al., 2012; Fung et al., 2010; Gkoulalas-Divanis and Loukides, 2011; LeFevre et al., 2006; Wang and Yang, 2011). In this context, collaborative in sharing healthcare information using HISs based on privacy preservation rarely handles healthcare information sharing among physicians and researchers at different places need to collaborate and communicate with each other to provide safer and more accessible to improve research findings that lead to enhanced care to patients. The need to address such collaboration among physicians and researchers in research activities based on privacy preservation is of utmost importance. A number of studies on the benefits of HISs have been conducted in the healthcare sector. These studies determined their effect on outcomes, including quality, efficiency, and provider satisfaction. Three systematic reviews of peer-reviewed studies about the benefits of adopting HISs in healthcare systems have been conducted and covered from 1994 to 2010 (Buntin and Burke, 2011; Goldzweig et al., 2009; Wu et al., 2006). Buntin and Burke (2011) cover the findings of these reviews and mentioned that 92% of recent articles on health IT reached conclusions that were generally positive (Buntin and Burke, 2011). Moreover, they found that the benefits of this technology were beginning to emerge in smaller practices and organizations as well as in large organizations that were early adopters. However, dissatisfaction with EMRs among some providers continued to hinder the potential of health IT. These realities highlight the need for studies that document the challenging aspects of the more strategic implementation of health IT and how these challenges may be addressed. Figure 1 summarizes the aforementioned findings on the benefits of health IT to the healthcare sector.

The collaboration among physicians in sharing information using HISs in the patient treatment or research activities within the hospital environment in many developing countries is very weak (Organization, 2010; Reddy et al., 2011). This weak occurs due to decentralized and autonomous units and lack of shared goals within healthcare systems; many HISs are isolated from one another because of the fragmented nature of healthcare systems (Fried et al., 2011). Disintegrated HISs and manual systems hinder information sharing and collaboration among physicians, thus impeding optimal use of healthcare resources and delaying because large amounts of data are difficult to manage and control in a

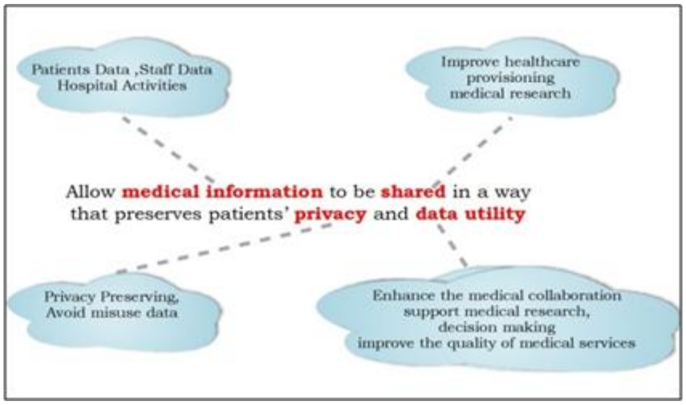
Patients Data ,Staff Data
Hospital Activities
Improve healthcare
provisioning
medical research
Allow medical information to be shared in a way that preserves patients' privacy and data utility
Privacy Preserving,
Avoid misuse data
Enhance the medical collaboration
support medical research,
decision making
improve the quality of medical services

Figure 2. Research motivation (Gkoulalas-Divanis and Loukides, 2011).

system that uses paper (Tierney et al., 2010; Van Vactor, 2012) introduced another important factor that affects collaboration among physicians, that is, privacy concerns raise the necessity of improving collaboration among medical staff through HISs. Effective implementation of HISs requires trust from both the providers who use them and the patients they serve (Blumenthal, 2009;Chen et al., 2012; Goldzweig et al., 2009). In such cases, sharing information regarding patients' treatment and medical researches among hospitals is difficult. The aforementioned factors critically affect technology acceptance in hospitals and collaboration among physicians, which can lead to poor patient outcomes (Reddy et al., 2011). The bigger challenge is strengthening sharing of healthcare information among physicians and researchers in same or different hospital, many of which still rely on paper-based records. As such, introducing new activities to hospitals is a difficult process. These activities are important in enhancing healthcare services. Collaborative HISs based on privacy preservation rarely handles healthcare information sharing among physicians and researchers at different places need to collaborate and communicate with each other to provide safer and more accessible to improve research findings that lead to enhanced care to patients. The need to address such collaboration among physicians and researchers in research activities based on privacy preservation is of utmost importance.

The privacy preservation is an important issue when dealing with personal data and can be considered as the backbone for the sharing data process. There are numerous real-world applications which require sharing data while meeting specific privacy constraints. Consequently, the literature review in this section aims to clarify the privacy preservation data sharing challenges.

The recent studies refer to the increase privacy and security consciousness has lead to increased research and development of methods that compute useful information in a secure fashion (Clifton et al., 2004; Fung et al., 2010). Data sharing have been a long standing challenge for the database community. This need has become critical in numerous contexts, including integrating data on the Web and at enterprises, building ecommerce market places, sharing data for scientific research, data exchange at government agencies, monitoring health crises, and improving homeland security (Clifton et al., 2004). Additional to large amounts of personal health data are being collected and made available through existing and emerging technological media and tools. While use of these data has significant potential to facilitate research, improve quality of care for individuals and populations, and reduce healthcare costs, many policy-related issues must be addressed before their full value can be realized. These include the need for widely agreed-on data stewardship principles and effective approaches to reduce or eliminate data silos and protect patient privacy (Hripcsak et al., 2014).

Unfortunately, data integration and sharing are hampered by legitimate and widespread privacy concerns (Clifton et al., 2004; Fung et al., 2010). Companies could share information to boost productivity, but are prevented by fear of being exploited by competitors or antitrust concerns. Sharing healthcare data could improve scientific research, but the cost of obtaining consent to use individually identifiable information can be prohibitive and these efforts must engage patients as partners (Hripcsak, et al., 2014). Sharing healthcare and consumer data enables early detection of disease outbreak (Tsui et al., 2003), but without provable privacy protection it is difficult to extend these surveillance measures nationally or internationally. Besides effective public safety and health care, collaboration and sharing between public agencies, and public and private organizations, can have a strong positive impact on public safety.

The continued exponential growth of distributed personal data could further fuel data integration and sharing applications, but may also be stymied by a privacy backlash. It is critical to develop techniques to enable the integration and sharing of data without losing privacy. As noted above, there is widespread agreement on the value of personal health data for many uses beyond direct patient care and treatment. Thus, discussions about the privacy preservation data sharing are more important than ever. As part of the overall problem, the literature review in this study aims to cover the privacy preserving data sharing as mentioned in the recent studies. The recent studies indicate to the emergent privacy issues of healthcare data are important issue. According to Gkoulalas and Loukides (2011) mentioned that 62% of individuals worry that their electronic medical records will not remain confidential (Gkoulalas-Divanis and Loukides, 2011), and 35% expressed privacy concerns regarding the collaboration (publishing and sharing) of their data (Ludman et al., 2010), Figure 2 shows the motivation for this work.

The literature review in this study aims to cover the privacy preserving data sharing as mentioned in the recent studies, in order to improve the collaboration among medical staff (relation management) with regard to medical data sharing for research through review and classification methods of privacy protection. The recent studies indicate to the emergent privacy issues of healthcare data are important issue. In the sections that follow, we briefly explain the related works and highlight related literature, collaboration in sharing healthcare information based on privacy preservation (relation between sharing and privacy), state of the art privacy preserving, privacy preservation and technical contribution, privacy preservation models, and proposed model to sharing healthcare information based on control privacy preservation.

RELATED WORKS

Privacy protection is an important issue particularly with regards to personal data that must have stringent policies on sharing. A definition on privacy protection has specified that access to published data should not allow potential attackers to learn anything beyond what target victims had permitted to disclose, which is in contrast to having no access to the database or the background knowledge of the potential attacker that he has obtained from other sources (Dalenius, 1977). The development of information technology and the collection of electronic information by data owners, such as governments, corporations, and individuals, have facilitated higher instances of data sharing and knowledge management. Driven by mutual benefits, these data owners have created broad opportunities for knowledge management and for information retrieval. Recent developments have helped improve decision making, particularly in the fields of medical information, research, and public health organization, among others. Many approaches have been proposed for different data publishing needs in different fields. Data sharing requires control and management to ensure system integration. Integration is required specifically in the management of patient data to secure sensitive information such as the identity of the patients (Gkoulalas-Divanis and Verykiosc, 2009; Qi and Zong, 2012). Several studies had focused on the management of data, such as in medical applications, to ensure system integration. However, management and sharing of data in different fields can lead to misuse of information, disclosure of the identification of the data owner, and other related problems (Clifton et al., 2004; Rashid et al., 2012). The primary goal in privacy preservation is the protection of sensitive data before they are released for analysis or for re-publication. Data may be kept at centralized or at distributed data storage areas. In this scenario, appropriate algorithms or techniques should be used to protect any sensitive information during the knowledge discovery process. Many approaches can be adopted for privacy-preserving data mining (Kaye et al., 2010).

An important aspect on privacy-preserving data mining algorithms and on tools for development and evaluation is to select the appropriate evaluation criteria. The reality, however, is that privacy-protected data mining algorithms with a variety of indicators are not better than other algorithms. Generally, an algorithm may be practical in terms of performance or may be slightly better than others. Users must be provided with a set of metrics to enable them to choose the best appropriate algorithms for data privacy preservation. Subsequently, we formulated a simple introduction on algorithm performance, data utility, privacy protection degree, and on the difficulty of different data mining techniques (Qi and Zong, 2012). In algorithm performance, the algorithm with O(n2) complexity polynomial time is more efficient than those with O(en) index of complexity. An alternative approach is necessary to evaluate time requirements in terms of average number of operations to reduce the frequency of sensitive information appearing below a specified threshold. Possibly, this value does not provide an absolute measure, but it can be capable of performing a fast comparison among different algorithms (Qi and Zong, 2012). Data utility is a very important issue in the implementation of data privacy protection. To hide sensitive information, false information may be inserted into the database or data values can be blocked. Although sample techniques do not modify the information stored in the database, they can exhibit a reduction because of the presentation of incomplete information (Qi and Zong, 2012). In the degree of privacy protection, the privacy protection policy prevents the downgrade of information to a certain threshold, though hidden information can be derived by some uncertainty. The uncertainty reconstructed by hidden information can evaluate the sanitation algorithm. A solution can set a maximum on perturbation information from the execution perspective, and then consider achieving the degree of uncertainty by measuring the constraints of different purification methods. We intend to define an algorithm that can achieve the highest uncertainty and that is better than all other algorithms (Qi and Zong, 2012). In difficulty of different data mining techniques, we must measure the difficulty of data mining algorithms, which differ from the purification method, to provide full estimation on the purification method called parameter horizontal difficulty. Parameter estimation must consider the data mining classification, which is important to the test. Alternatively, we may need to develop a formal framework that can ensure privacy assurance for an entire class of sanitization algorithms upon testing one against pre-selected data sets (Qi and Zong, 2012).

The recent studies refer to the increase privacy and security consciousness has lead to increased research and development of methods that compute useful information in a secure fashion (Clifton et al., 2004; Fung et al., 2010). Data sharing have been a long standing challenge for the database community. In other words, great concern has been directed on the control of data and it’s sharing to make it available to their owners. Some reviewers and researchers have even suggested the use of covert techniques which isolate data such as encryption technology. Different ways of protecting data have been dealt with in recent research. The methods previously introduced include information on how to spread and use data in research, decision making, scientific analyses, and other purposes (Fung et al., 2010). First, the concern is how to control data sharing and management and avoid the risk of publishing data that may lead to revealing the real data. Second, there is lack of unity among the collected data, and their sources vary as they are collected from various points such as governments, hospitals, companies, and so on. Third, the data collected may contain errors. How data are processed and formatted before access requires a high level of analysis techniques to extract and determine knowledge and relationships hidden.

To identify the relationships among different data and their influence on the results, they must be accurate and correct, as one type of data relies on the results of the analysis. Examples are the reasons for the spread of a particular disease in a particular area in the medical field, the losses incurred by a company after a change in business strategy, and the low standards of living in a society. The main objective of the present research is to control management and sharing of data in the medical field, which mainly involves "patient data". The main objectives of the present research is to sharing healthcare information based on privacy preservation and keep data utility for secondary purposes such as research.

COLLABORATION HEALTHCARE INFORMATION BASED ON PRIVACY PRESERVATION

Recently, many healthcare organizations are adopting Customer relationship management (CRM) as a strategy, which involves using technology to organize, automate, and coordinate business processes, in managing interactions with their patients. CRM with the Web technology provides healthcare providers the ability to broaden their services beyond usual practices, and thus offers suitable environment using latest technology to achieve superb patient care (Anshari and Almunawar, 2012).

There are two basic types of healthcare CRMs, one is for a healthcare organization to stay in contact with their patients, and the other is for a healthcare organization to stay in contact with referring organizations. In other hand, privacy is critical factor when patients’ information used in other treatment purposes (Fung et al., 2010; Gkoulalas-Divanis and Loukides, 2011).

One of the most interesting aspects in medical care is how to manage the relationship between healthcare providers and patients (Anshari and Almunawar, 2012). Fostering relationship leads to maintain loyal customer, greater mutual understanding, trust, patient satisfaction, and patient involvement in decision making (Glanz et al., 2008). Furthermore, effective communication is often associated with improved physical health, more effective chronic disease management, and better health related quality of life (Arora, 2003). On the other hand, failure in managing the relationship will affect to the patient dissatisfaction, distrust towards systems, patient feels alienated in the hospital, and jeopardize business survivability in the future.

In this context, Usually, CRM is applied in the business field but not in the medical one. The application of the CRM model can result in desirable results through collaboration among hospital in patients treatment and other purposes such as data analysis, research. In other hand, Data mining has been used intensively and extensively by many organizations (Anshari and Almunawar, 2012). In healthcare, data mining is becoming increasingly popular, if not increasingly essential. Data mining applications can greatly benefit all parties involved in the healthcare industry. For example, data mining can help healthcare insurers detect fraud and abuse, healthcare organizations make customer relationship management decisions, physicians identify effective treatments and best practices, and patients receive better and more affordable healthcare services The huge amounts of data generated by healthcare transactions are too complex and voluminous to be processed and analyzed by traditional methods. Data mining provides the methodology and technology to transform these mounds of data into useful information for decision making (Koh and Tan, 2011). In healthcare, data mining is becoming increasingly popular, if not increasingly essential. Several factors have motivated the use of data mining applications in healthcare. The existence of medical insurance fraud and abuse, for example, has led many healthcare insurers to attempt to reduce their losses by using data mining tools to help them find and track offenders (Anshari and Almunawar, 2012; Christy, 1997). Fraud detection using data mining applications is prevalent in the commercial world, for example, in the detection of fraudulent credit card transactions. Recently, there have been reports of successful data mining applications in healthcare fraud and abuse detection (Milley, 2000). Another factor is that the huge amounts of data generated by healthcare transactions are too complex and voluminous to be processed and analyzed by traditional methods. Data mining can improve decision-making by discovering patterns and trends in large amounts of complex data (Biafore, 1999). Insights gained from data mining can influence cost, revenue, and operating efficiency while maintaining a high level of care (Silver et al., 2001).

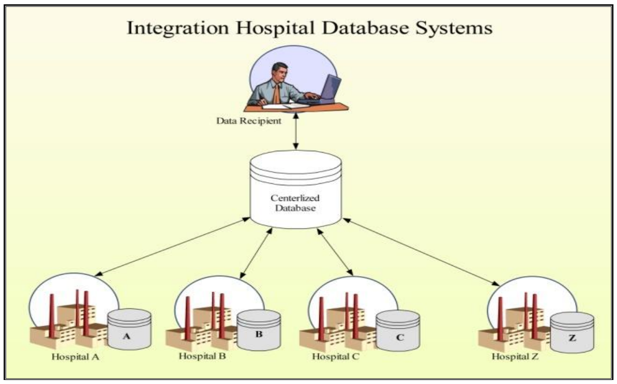

Figure 3. Integration hospital database system.

Healthcare organizations that perform data mining are better positioned to meet their long-term needs, Benko giving an illustration of a healthcare data mining application; and finally, highlighting the limitations of data mining and offering some future directions Cios and Moore31 have argued that data problems in healthcare are the result of the volume, complexity and heterogeneity of medical data and their poor mathematical characterization and non-canonical form. Further, there may be ethical, legal and social issues, such as data ownership and privacy issues, related to healthcare data. The quality of data mining results and applications depends on the quality of data (Koh and Tan, 2011).

Recent studies have shown that the development of effective collaborative HISs to support collaborative work among medical staff, especially among physicians and researchers, requires the use of real data. This result is based on the fact that the collaborative HIS approach requires appropriate, flexible, and comprehensive healthcare information based on user (Kuziemsky et al., 2012; Kuziemsky and Varpio, 2011; Lezzar et al., 2012; Reddy et al., 2011; Ruxwana et al., 2010; Scandurra et al., 2008). The findings of the review here indicate strong relationship between collaboration in sharing healthcare information and privacy preservation as mentioned in recent studies, in order to development of effective collaborative HISs to support collaborative work and improve patients outcome. Many researchers in this area proposed healthcare system models for healthcare information sharing among medical staff, and few studies focused on the research on healthcare system and privacy preservation in health sector. However, such models are not flexible in structure and are difficult to manage and control because of the enormous data in complex healthcare systems. The Figure 3 shows the Integration HISs.

In the past few years, research communities have responded to the challenges of privacy preservation through collaborative activities in sharing data as mentioned in (Clifton and Atallah, 2007) to eliminate privacy concerns from patients and help medical institutions or participants comply with privacy protection regulations. These approaches encompass several fields of research. The problems they are trying to address could be classified into three categories:

The first category focuses on privacy protection in data sharing during data usage. These kinds of approaches attempt to protect patient privacy by transforming the healthcare data before they are shared. The privacy information may be wiped or reduced after the transforming process. The de-identification approach simply detects the private data and deletes them (Neamatullah et al., 2008). To retain the usability of the transformed data as much as possible, many new models and methods are proposed. Privacy-preserving data publishing models, such as K-anonymity and l-diversity (Fung et al., 2010), and privacy-preserving data mining models and methods, such as privacy-preserving decision trees and associate rule mining (Aggarwal and Philip, 2008), have been developed as a result of these studies. The second category focuses on privacy data management. Many access control models and systems have been developed to enhance the flexibility of privacy data management and compliance with regulations. Elements such as access purpose, data content, and personal preferences have been brought into these data access management models (Byun et al., 2005; Smith, 2001). The third category focuses on privacy data storage and management. Privacy for data storage and management in a cloud environment has attracted plenty of attention in recent years. Approaches for privacy-aware data storage and auditing in a cloud environment are proposed to protect private data (Itani et al., 2009; Wang et al., 2010).

All approaches listed above may be used in privacy data sharing or management in some way. Many abstract frameworks have been proposed to realize privacy protection during data sharing, such as a framework for privacy preserving data sharing proposed by Chen (2004). Kennelly (2009) developed an Internet data-sharing framework for balancing privacy and utility. However, to the best of our knowledge, few research works about healthcare data sharing frameworks that preserve the privacy of users offer a practical view for real life application (Chen et al., 2012).

However, one set of methods that would allow health information to be used and disclosed under existing legal frameworks is de-identification. De-identification refers to a set of methods that can be applied to data to ensure that the probability of assigning a correct identity to a record in the data is very low (El Emam and Fineberg, 2009; El Emam et al., 2011). Recent studies (Bayardo and Agrawal, 2005; Campan and Truta, 2009; El Emam et al., 2012; El Emam and Dankar, 2008; El Emam et al., 2009; Goryczka et al., 2011; Jiang and Clifton, 2006; Jurczyk and Xiong, 2009; LeFevre et al., 2005; Parmar et al., 2011; Sacharidis et al., 2010; Sokolova et al., 2012; Sweeney, 2002a, b; Tassa and Gudes, 2012; Truta and Vinay, 2006) indicate that the K-anonymity model provides a formal way of generalizing this concept because K-anonymity provides a measure of privacy protection by preventing the re-identification of data to fewer than a group of K data items. As stated in Sweeney and Samarati (Samarati, 2001;Sweeney, 2002a, b), a data record is K anonymous if and only if it is indistinguishable from its identifying information from at least K-specific records or entities. The key step in making data anonymous is to generalize a specific value. Generalized data can be beneficial in many situations as stated in (Chen et al., 2012; Jiang and Clifton, 2006). Many applications are used to generalize data in a many areas, including medical research, education studies, and targeted marketing.

STATE-OF-THE-ART PRIVACY PRESERVING

This study covers a review of the most relevant areas below and discuss how our work levels up with recent state-of-the-art systems.

Privacy preservation in data publication

The preservation of privacy when publishing data for centralized databases has been examined intensively in recent years. One thread of work aims at devising privacy principles such as k-anonymity and subsequent principles that address problems, which in turn serve as criteria for judging whether a published data set enables privacy protection (Nergiz and Clifton, 2007; Sweeney, 2002b). Another body of work has contributed to the development of an algorithm that transforms a data set to meet one of the privacy principles (dominantly k-anonymity). However, most of these works have focused only on structured data (Gardner and Xiong, 2009; Li et al., 2007; Xiao and Tao, 2007).

Medical text de-identification

In the medical informatics community, there have been efforts in de-identifying medical text documents (Gardner and Xiong, 2009; Sweeney, 2002b; Zhong et al., 2005). Most of them use a two-step approach which extracts the identifying characters first and then removes or masks the attributes for de-identification purposes. Most of them are specialized for specific document types, for example, pathology reports only (Gardner and Xiong, 2008; Zhong et al., 2005). Some systems focus on a subset of Health Insurance Portability and Accountability Act (HIPAA) identifiers, for example, name only (Aramaki et al., 2006; Gardner and Xiong, 2009), whereas others focus on differentiating protected health information (PHI) from non-PHI (Gardner and Xiong, 2009). Most importantly, most of these studies rely on simple identifier removal or grouping techniques, and they do not take advantage of recent research developments that guarantee a more formalized notion of privacy while increasing data utility.

Information extraction

Extracting atomic identifiers and sensitive characters (such as name, address, and disease) from unstructured text such as pathology reports can be seen as an application of the named entity recognition (NER) problem (Neumann, 2010). NER systems can be roughly classified into two categories, both of which are applied in medical domains for de-identification. The first uses grammar-based or rule-based techniques (Gardner and Xiong, 2008). Unfortunately, such hand-crafted systems may take months of work by experienced domain experts, and the rules will likely change for different data repositories. The second category uses statistical learning approaches such as support vector machine (SVM)-based classification methods. However, an SVM-based method such as that introduced by Sibanda and Unuzer (Sibanda and Uzuner, 2006) only performs binary classification of the terms into PHI or non-PHI. It does not also allow statistical de-identification which requires knowledge on different types of identifying characters.

PRIVACY PRESERVATION AND TECHNICAL CONTRIBUTION

In the following, the researcher explains technical contributions of the survey to data privacy through the control and sharing of data in knowledge management. We focus on six aspects of technical contributions, which we consider to be the most interesting (Xiao, 2009).

Personalized privacy preservation

We examined the work of (Xiao and Tao, 2006) on the publication of sensitive data using generalization, the most popular anonymization methodology in the literature. The existing privacy model for generalized tables (that is, noisy microdata obtained through generalization) exerts the same amount of protection on all individuals in the data set without catering to their concrete needs. For example, in a set of medical records, a patient who has contracted flu would receive the same degree of privacy protection as a patient suffering from cancer, despite the willingness of the former to reveal his/her symptoms directly (mainly because flu is a common disease) (Xiao and Tao, 2006). Motivated by this, we propose a personalized framework that allows each individual to specify his/ her preferred privacy protection in relation to his/her data. Based on this framework, we devised the first privacy model that considers personalized privacy requests. We also developed an efficient algorithm for computing generalized tables that conform to the model. Through extensive experiments, we show that our solution outperforms other generalization techniques by providing superior privacy while incurring the least possible information loss (Xiao and Tao, 2006).

Republishing dynamic data sets

Data collection is often a continuous process, where tuples are inserted into and deleted from the microdata as time evolves. Therefore, a data publisher may need to republish the microdata at multiple times to reflect the most recent changes. Such republication is not supported by conventional generalization techniques because microdata are assumed to be static (Xiao and Tao, 2007). We address this issue by proposing an innovative privacy model called m-invariance which secures the privacy of any individual involved in the republication process, even against a rival who exploits the correlations between multiple releases of the microdata. The model is accompanied by a generalization algorithm whose space and time complexity are independent of the number n of generalized tables that have been released by the publisher. This property of the algorithm is essential in the republication scenario, where n increases monotonically with time (Xiao and Tao, 2007).

Complexity of data anonymization

We have presented the first study on the complexity of producing generalized tables, which conform to ℓ-diversity, the most commonly adopted privacy model. We note that achieving ℓ-diversity with minimum information loss is NP-hard for any ℓ larger than two and any data set that contains at least three distinct sensitive values. Considering this, we developed an O(ℓ.d)-approximation algorithm, where d is the number of QI characters contained in the microdata (Xiao, 2008). Aside from its theoretical guarantee, the proposed algorithm works fairly well in practice and considerably outperforms state-of-the-art techniques in several aspects (Xiao, 2008).

Transparent anonymization

Previous solutions for data publication consider the idea that the rival controls certain prior knowledge about each individual. However, they overlook the possibility that the rival may also know the anonymization algorithm adopted by the data publisher. Thus, an attacker can compromise the privacy protection enforced by the solutions by exploiting various characteristics of the anonymization approach (Xiao, 2008). To address this problem, we propose the first analytical model for evaluating the disclosure risks in generalized tables under the assumption that everything involved in the anonymization process, except the data set, is public knowledge. Based on this model, we developed three generalization algorithms to ensure privacy protection, even against a rival who has a thorough understanding of the algorithms. Compared with state-of-the-art generalization techniques, our algorithms not only provide a higher degree of privacy protection but also satisfactory performance in terms of information distortion and overhead estimation (Xiao, 2008).

Anonymization via anatomy

While most previous work adopts generalization to anonymize data, we propose a novel anonymization method anatomy which provides almost the same privacy guarantee as generalization does. However, it significantly outperforms it in terms of the accuracy of data analysis on the distorted microdata (Xiao and Tao, 2006). We provide theoretical justifications for the superiority of anatomy over generalization and develop a linear time algorithm for anonymizing data via anatomy. The efficiency of our solution was verified through extensive experiments.

Dynamic anonymization

We propose dynamic anonymization which produces a tailor-made anonymized version of the data set for each query given by users; the anonymized data increases the accuracy of the query result. Privacy preservation is achieved by ensuring that no private information is revealed despite combining all anonymized data (Xiao, 2008). For example, even if the rival obtains every anonymized version of the data set, he/she would not be able to infer the sensitive value of any individual. Through extensive experiments, we show that compared with existing techniques, dynamic anonymization significantly improves the accuracy of queries on the anonymized data (Xiao, 2008).

PRIVACY PRESERVATION MODELS

Recent developments in healthcare technology enable the collection, storage, management, and sharing of massive amounts of medical data (Lau et al., 2011). HISs are increasingly adopted in the healthcare sector (Dean et al., 2010; Makoul et al., 2001). The use of HISs allows specialists to access comprehensive medical information,

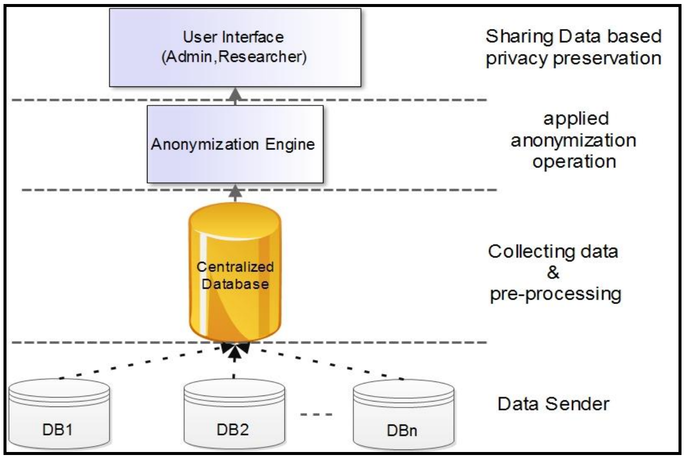
User Interface
(Admin,Researcher)
Sharing Data based privacy preservation
Anonymization Engine
applied anonymization operation
Centralized Database
Collecting data & pre-processing
DB1
DB2
DBn
Data Sender

Figure 4. Collaborative healthcare information management system based on privacy preservation.

to extract knowledge and reduce medical errors, as well as to collaborate with other specialists and healthcare entities to improve the diagnosis and treatment of diseases. At the same time, reusing medical data offers the potential to improve medical research findings. However, reusing medical data must be performed in a way that addresses important privacy concerns.

Preserving the privacy of medical data is not only an ethical but also a legal requirement that is posed by several data sharing regulations and policies worldwide. For example, in 1996, the Health Insurance Portability and Accountability Act (HIPAA) title II was enacted in the USA (Act, 1996; Nosowsky and Giordano, 2006). One of the purposes of this act is to increase the protection of patients' medical records against unauthorized usage and disclosure. Hospitals, clinical offices, health insurance companies, and other entities governed by HIPAA are asked to comply with regulations. In 1997, the European Council announced Recommendation R (97) 5 regarding the protection of medical data to enhance the protection of personal healthcare data (DIRECTIVE, 1997). Similar regulations have been enacted in many other countries (Chen et al., 2012). For example, contracts and agreements cannot guarantee that sensitive data will not be carelessly misplaced and end up in the wrong hands. A task of the utmost importance is developing methods and tools for publishing data in a more hostile environment, so that the published data (shared data) remains practically useful while preserving individual privacy. This undertaking is termed privacy-preserving data publishing (Fung et al., 2009; Gkoulalas-Divanis and Loukides, 2011; Gkoulalas-Divanis and Verykiosc, 2009). Privacy-preserving data publishing and information security communities have recently begun addressing these issues. Numerous techniques have been developed to address the first problem, which is avoiding potential misuse posed by an integrated data warehouse (Vaidya et al., 2006). Many abstract frameworks have been proposed to realize privacy protection during data sharing, such as a framework for privacy preserving data sharing proposed by Chen (2004). Kennelly (2009) developed an Internet data-sharing framework for balancing privacy and utility. However, to the best of our knowledge, few research works about healthcare data sharing frameworks that preserve the privacy of users offer a practical view for real life application (Chen et al., 2012).

The finding form this section indicates to K-anonymity model is suitable methods in sharing information in healthcare sector. The main features of the K-anonymity model as mentioned in recent literature: K-anonymity is a simple and effective (Sweeney, 1997, 2002b) model that provides a measure of privacy protection by preventing the re-identification of data to fewer than a group of K data items (Jiang and Clifton, 2006; Narayanan and Shmatikov, 2009), providing a formal way of generalizing this concept (Samarati, 2001; Sweeney, 2002a, b), and minimizing data utility loss while limiting disclosure risk to an acceptable level (Morton et al., 2012). In addition, the K-anonymity model is a simple and practical model for data privacy preservation (Chiu and Tsai, 2007), and it guarantees that the data released are accurate (Barak et al., 2007).

COLLABORATIVE HEALTHCARE INFORMATION SYSTEM: PROPOSED MODEL

The collaborative healthcare information management system, which was based on the k-anonymization model and generalization technique, was developed to achieve the objective of improving collaboration and outcomes based on a privacy preservation approach. The proposed framework comprises four phases. The first phase involves collecting data from different HISs, and then sending the data to a central database. The second phase involves data pre-processing, such as missing values, inconsistent data, data integration, data selection, and data transformation. The third phase involves processing data based on the anonymization engine, which applies the anonymization operation based on the data generalization technique; this phase involves "a strategy for protecting individual privacy in released microdata records". The fourth phase involves sharing data among researchers based on privacy preservation as shown in Figure 4.

The idea is that by reconstructing a more "general" and semantically consistent domain for the attributes and transforming its values to this domain, identifying individuals by linking this attribute with external data would be much more difficult. From the perspective of information communication technology (ICT), the CHIMS construction was developed on the basis of an agent-based technique for linking the CHIMS units in different departments at hospitals using Web-based application tools; in this stage collecting healthcare data from different HISs departments, and then sending the data to a central database. The second stage pre-processing data in this study the researcher assume the collected data of hospital departments is clear. Stage three collected healthcare data send to anonymization engine in order to privacy preservation; to anonymize data was applied generalization, which transforms attribute values of non-sensitive attributes in the data into values ranges, so as to prevent an adversary from identifying individuals by linking these attributes with public available information. In hospital environment the collaboration among medical staff increases the awareness of team members regarding their respective knowledge and skills, which leads to further improvements in decision making and improve the research findings in healthcare sector. Consequently, Collaboration is an important requirement in health information systems (HISs) because it produces reliable and rigorous evidence that can inform critical decisions related to healthcare services. It aids in the provision of proper, fast treatment to patients, and healthcare information for research.

CONCLUSION

Collaboration in HISs is important in providing proper and fast treatment to patients and suitable medical data for research. Collaboration among current healthcare departments is important in addressing most HISs problems and in satisfying all system requirements. These requirements must maximize information flows and storage among HISs units to provide information in an appropriate and timely manner based on privacy preservation. Anonymization approach has been successfully used to provide privacy preservation and to maintain data utility. Therefore, this study improved the collaboration research among physicians and researchers by developing CHIMS based on the k-anonymization model, which in turn addressed privacy preservation and improved healthcare services through adoption in HISs.

Conflict of Interest

The authors have not declared any conflict of interests.

REFERENCES

Act A (1996). Health insurance portability and accountability act of 1996. Public Law, 104, 191.

Aggarwal CC, Philip SY (2008). A general survey of privacy-preserving data mining models and algorithms: Springer.

Aggelidis VP, Chatzoglou PD (2009). Using a modified technology acceptance model in hospitals. International J. Med. Informatics. 78(2):115.

Ahmed NS, Yasin NM (2012). Improvement the Cooperation Feature in Distributed Healthcare Information Systems Based on the Fractal Approach: An Empirical Study. Adv. Mater. Res. 463:861-867.

Ammenwerth E, Gräber S, Herrmann G, Bürkle T, König J (2003). Evaluation of health information systems—problems and challenges. Int. J. Med. Inf. 71(2):125-135.

Anshari M, Almunawar MN (2012). Evaluating CRM implementation in healthcare organization. arXiv preprint arXiv: 1204:3689.

Aramaki E, Imai T, Miyo K, Ohe K (2006). Automatic deidentification by using sentence features and label consistency.

Arora NK (2003). Interacting with cancer patients: the significance of physicians' communication behavior. Soc. Sci. Med. 57(5):791-806.

Barak B, Chaudhuri K, Dwork C, Kale S, McSherry F, Talwar K (2007). Privacy, accuracy, and consistency too: a holistic solution to contingency table release. Paper presented at the Proceedings of the twenty-sixth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems.

Bayardo RJ, Agrawal R (2005). Data privacy through optimal k-anonymization.

Biafore S (1999). Predictive solutions bring more power to decision makers. Health Manage. Technol. 20(10):12.

Blumenthal D (2009). Stimulating the adoption of health information technology. New Engl. J. Med. 360(15):1477-1479.

Buntin MB, Burke MF, Hoaglin MC, Blumenthal D (2011). The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Affairs, 30(3):464-471.

Byun J-W, Bertino E, Li N (2005). Purpose based access control of complex data for privacy protection. Paper presented at the Proceedings of the tenth ACM symposium on Access control models and technologies.

Campan A, Truta T (2009). Data and structural k-anonymity in social networks. Privacy, Security, and Trust in KDD. pp. 33-54.

Chen L, Yang J-J, Wang Q, Niu Y (2012). A framework for privacy-preserving healthcare data sharing. Paper presented at the e-Health Networking, Applications and Services (Healthcom), 2012 IEEE 14th International Conference on.

Chen L, Yang JJ, Wang Q, Niu Y (2012). A framework for privacy-preserving healthcare data sharing. Paper presented at the e-Health Networking, Applications and Services (Healthcom), 2012 IEEE 14th International Conference on; 01/2012.

Chiu C-C, Tsai C-Y (2007). A k-anonymity clustering method for effective data privacy preservation. Advanced Data Mining and Applications. pp. 89-99.

Christy T (1997). Analytical tools help health firms fight fraud. Insurance and Technology, 22(3):22-26.

Clifton C, Atallah M (2007). Collaborative Research: ITR: Distributed Data Mining to Protect Information Privacy.

Clifton C, Kantarcioğlu M, Doan A, Schadow G, Vaidya J, Elmagarmid A, Suciu D (2004). Privacy-preserving data integration and sharing. Paper presented at the Proceedings of the 9th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery.

Dalenius T (1977). Towards a methodology for statistical disclosure control. Statistik Tidskrift, 15(429-444):2-1.

Dean BB, Lam J, Natoli JL, Butler Q, Aguilar D, Nordyke RJ (2010). Use of Electronic Medical Records for Health Outcomes Research: A Literature Review. Med. Care Res. Rev. 66(6):611-638.

DIRECTIVE HAT (1997). Council Directive 97/43/Euratom of 30 June 1997 on health protection of individuals against the dangers of ionizing radiation in relation to medical exposure, and repealing Directive 84/466/Euratom. Official J. No. L, 180(09/07):0022-0027.

El Emam K, Arbuckle L, Koru G, Eze B, Gaudette L, Neri E, Gluck J (2012). De-identification methods for open health data: the case of the Heritage Health Prize claims dataset. J. Med. Internet Res.14(1).

El Emam K, Dankar FK (2008). Protecting privacy using k-anonymity. J. Am. Med. Inf. Assoc. 15(5):627-637.

El Emam K, Dankar FK, Issa R, Jonker E, Amyot D, Cogo E, Vaillancourt R (2009). A globally optimal k-anonymity method for the de-identification of health data. J. Am. Med. Inf. Assoc. 16(5):670-682.

El Emam K, Fineberg A (2009). An overview of techniques for de-identifying personal health information. Access to Information and Privacy Division of Health Canada.

El Emam K, Jonker E, Fineberg A (2011). The Case for De-identifying Personal Health Information.

Ernstmann N, Ommen O, Neumann M, Hammer A, Voltz R, Pfaff H (2009). Primary care physician's attitude towards the GERMAN e-Health Card Project—Determinants and Implications. J. Med. Syst. 33(3):181-188.

Fried B, Carpenter WR, Deming WE (2011). Understanding and improving team effectiveness in quality improvement. McLaughlin and Kaluzny's Continuous Quality Improvement in Health Care, P. 117.

Fung B, Wang K, Wang L, Hung PCK (2009). Privacy-preserving data publishing for cluster analysis. Data and Knowledge Engineering, 68(6):552-575.

Fung BCM, Wang K, Chen R, Yu PS (2010). Privacy-preserving data publishing: A survey of recent developments. ACM Comput. Surv. 42(4):1-53. doi: 10.1145/1749603.1749605

Fung BCM, Wang K, Chen R, Yu PS (2010). Privacy-Preserving Data Publishing: A Survey of Recent Developments. Acm Computing Surveys, 42(4). doi: 1410.1145/1749603.1749605

Gaboury I, Bujold M, Boon H, Moher D (2009). Interprofessional collaboration within Canadian integrative healthcare clinics: Key components. Soc. Sci. Med. 69(5):707-715.

Gardner J, Xiong L (2008). HIDE: An integrated system for health information de-identification.

Gardner J, Xiong L (2009). An integrated framework for de-identifying unstructured medical data. Data and Knowledge Engineering, 68(12):1441-1451.

Gkoulalas-Divanis A, Loukides G (2011). Medical Data Sharing: Privacy Challenges and Solutions.

Gkoulalas-Divanis A, Verykiosc VS (2009). An overview of privacy preserving data mining. Crossroads, 15(4):6.

Glanz K, Rimer BK, Viswanath K (2008). Health behavior and health education: theory, research, and practice: John Wiley and Sons.

Goldzweig CL, Towfigh A, Maglione M, Shekelle PG (2009). Costs and benefits of health information technology: new trends from the literature. Health Affairs 28(2):w282-w293.

Goryczka S, Xiong L, Fung BC (2013). Secure Distributed Data Anonymization and Integration with m-Privacy.

Hillestad R, Bigelow J, Bower A, Girosi F, Meili R, Scoville R, Taylor R (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. Health Affairs, 24(5):1103-1117.

Hripcsak G, Bloomrosen M, FlatelyBrennan P, Chute CG, Cimino J, Detmer DE, Hammond WE (2014). Health data use, stewardship, and governance: ongoing gaps and challenges: a report from AMIA's 2012 Health Policy Meeting. J. Am. Med. Inf. Assoc. 21(2):204-211.

Itani W, Kayssi A, Chehab A (2009). Privacy as a service: Privacy-aware data storage and processing in cloud computing architectures. Paper presented at the Dependable, Autonomic and Secure Computing, DASC'09. Eighth IEEE International Conference on.

Jiang W, Clifton C (2006). A secure distributed framework for achieving k-anonymity. The VLDB Journal—The International Journal on Very Large Data Bases, 15(4):316-333.

Jiang, W., and Clifton, C. (2006). A secure distributed framework for achieving k-anonymity. Vldb Journal, 15(4), 316-333. doi: 10.1007/s00778-006-0008-z

Jurczyk P, Xiong L (2009). Distributed anonymization: Achieving privacy for both data subjects and data providers Data and Applications Security XXIII (pp. 191-207): Springer.

Kaye R, Kokia E, Shalev V, Idar D, Chinitz D (2010). Barriers and success factors in health information technology: A practitioner's perspective. J. Manage. Market. Healthcare 3(2):163-175.

Koh HC, Tan G (2011). Data mining applications in healthcare. J. Healthcare Inf. Manage. 19(2):65.

Kohn LT, Corrigan J, Donaldson MS (1999). To err is human: Building a safer health system. Committee on Health Care in America. Institute of Medicine: Washington (DC): National Academy Press.

Kuziemsky CE, O'Sullivan TL, Corneil W (2012). An Upstream-Downstream Approach for Disaster Management Information Systems Design. Paper presented at the Proceedings of the ISCRAM Conference.

Kuziemsky CE, Varpio L (2011). A model of awareness to enhance our understanding of interprofessional collaborative care delivery and health information system design to support it. Int. J. Med. Inf. 80(8):e150-e160.

Lau EC, Mowat FS, Kelsh MA, Legg JC, Engel-Nitz NM, Watson HN, Whyte JL (2011). Use of electronic medical records (EMR) for oncology outcomes research: assessing the comparability of EMR information to patient registry and health claims data. Clin. Epidemiol. 3:259.

LeFevre K, DeWitt DJ, Ramakrishnan R (2005). Incognito: Efficient full-domain k-anonymity.

LeFevre K, DeWitt DJ, Ramakrishnan R (2006). Mondrian multidimensional k-anonymity.

Lezzar F, Zidani A, Atef C (2012). A Collaborative Web-based Application for Health Care Tasks Planning. Paper presented at the ICWIT.

Lu YC, Xiao Y, Sears A, Jacko JA (2005). A review and a framework of handheld computer adoption in healthcare. Int. J. Med. Inf. 74(5):409.

Ludman EJ, Fullerton SM, Spangler L, Trinidad SB, Fujii MM, Jarvik GP, Burke W (2010). Glad you asked: participants' opinions of re-consent for dbGaP data submission. J. Empirical Res. Human Res. ethics: JERHRE, 5(3):9.

Makoul G, Curry RH, Tang PC (2001). The use of electronic medical records communication patterns in outpatient encounters. J. Am. Med. Inf. Assoc. 8(6):610-615.

Milley A (2000). Healthcare and Data Mining Using data for clinical, customer service and financial results. Health Manage. Technol. 21(8):44-45.

Morton S, Mahoui M, Gibson PJ (2012). Data anonymization using an improved utility measurement. Paper presented at the Proceedings of the 2nd ACM SIGHIT symposium on International health informatics.

Narayanan A, Shmatikov V (2009). De-anonymizing social networks. Paper presented at the Security and Privacy, 2009 30th IEEE Symposium on.

Neamatullah I, Douglass MM, Li-wei HL, Reisner A, Villarroel M, Long WJ, Clifford GD (2008). Automated de-identification of free-text medical records. Bmc Medical Informatics and Decision Making, 8(1):32.

Nergiz ME, Clifton C (2007). Thoughts on k-anonymization. Data and Knowledge Engineering, 63(3):622-645.

Neumann RG (2010). Information Extraction. Architect. 2: 05.11.

Nosowsky R, Giordano TJ (2006). The Health Insurance Portability and Accountability Act of 1996 (HIPAA) privacy rule: implications for clinical research. Annu. Rev. Med. 57:575-590.

Organization WH (2010). Country Cooperation Strategy for WHO and Egypt 2010–2014. Cario. http://www. who. int/countryfocus/cooperation_strategy/ccs_egy_en. pdf.

Parmar AA, Rao UP, Patel DR (2011). Blocking based approach for classification Rule hiding to Preserve the Privacy in Database.

Qi X, Zong M (2012). An Overview of Privacy Preserving Data Mining. Procedia Environ. Sci. 12:1341-1347.

Rashid AH, Yasin NBM (2012). Anonymization Approach for Protect Privacy of Medical Data and Knowledge Management. Medical Informatics, Prof. Shaul Mordechai (Ed.), ISBN: 978-953-51-0259-5, InTech, DOI: 10.5772/37190.

Reddy MC, Gorman P, Bardram J (2011). Special issue on supporting collaboration in healthcare settings: The role of informatics. Int. J. Med. Inf. 80(8):541-543.

Ruxwana NL, Herselman ME, Conradie DP (2010). ICT applications as e-health solutions in rural healthcare in the Eastern Cape Province of South Africa. Health Inf. Manage. J. 39(1):17-26.

Sacharidis D, Mouratidis K, Papadias D (2010). K-anonymity in the presence of external databases. Knowledge and Data Engineering, IEEE Transactions on, 22(3):392-403.

Samarati P (2001). Protecting respondents identities in microdata release. Knowledge and Data Engineering, IEEE Transactions on, 13(6):1010-1027.

Scandurra I, Hägglund M, Koch S (2008). From user needs to system specifications: multi-disciplinary thematic seminars as a collaborative design method for development of health information systems. J. Biomed. Inf. 41(4):557-569.

Scott RE (2007). e-Records in health—Preserving our future. Int. J. Med. Inf. 76(5):427-431.

Sibanda T, Uzuner O (2006). Role of Local Context in De-identification of Ungrammatical, Fragmented Text. North American Chapter of Association for Computational Linguistics/Human Language Technology (NAACL-HLT).

Silver M, Sakata T, Su H-C, Herman C, Dolins SB, O Shea MJ (2001). Case study: how to apply data mining techniques in a healthcare data warehouse. J. Healthcare Inf. Manage. 15(2):155-164.

Smith HE (2001). A Context-Based Access Control Model for HIPAA Privacy and Security Compliance. SANS Security Essentials. CISSP.

Sokolova M, El Emam K, Arbuckle L, Neri E, Rose S, Jonker E (2012). P2P Watch: Personal Health Information Detection in Peer-to-Peer File-Sharing Networks. J. Med. Internet Res. 14(4).

Sweeney L (1997). Guaranteeing anonymity when sharing medical data, the Datafly System. Paper presented at the Proceedings of the AMIA Annual Fall Symposium.

Sweeney L (2002a). Achieving k-anonymity privacy protection using generalization and suppression. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(05):571-588.

Sweeney L (2002b). k-anonymity: A model for protecting privacy. International Journal of Uncertainty Fuzziness and Knowledge Based Systems, 10(5):557-570.

Tassa T, Gudes E (2012). Secure distributed computation of anonymized views of shared databases. ACM Transactions on Database Systems (TODS), 37(2):11.

Tierney WM, Achieng M, Baker E, Bell A, Biondich P, Braitstein P, McKown B (2010). Experience implementing electronic health records in three East African countries. Stud Health Technol. Inf. 160(Pt 1):371-375.

Truta TM, Vinay B (2006). Privacy protection: p-sensitive k-anonymity property. Paper presented at the Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on.

Tsui FC, Espino JU, Dato VM, Gesteland PH, Hutman J, Wagner MM. (2003). Technical description of RODS: a real-time public health surveillance system. J. Am. Med. Inf. Assoc. 10(5):399-408.

Vaidya J, Zhu M, Clifton CW (2006). Privacy preserving data mining (Vol. 19): Springer-Verlag New York Inc.

Van Vactor JD (2012). Collaborative leadership model in the management of health care. J. Bus. Res. 65(4):555-561.

Wang B, Yang J (2011). The state of the art and tendency of privacy preserving data mining.

Wang C, Wang Q, Ren K, Lou W (2010). Privacy-preserving public auditing for data storage security in cloud computing. Paper presented at the INFOCOM, 2010 Proceedings IEEE.

Wang SJ, Middleton B, Prosser LA, Bardon CG, Spurr CD, Carchidi PJ, Sussman AJ (2003). A cost-benefit analysis of electronic medical records in primary care. The Am. J. Med. 114(5):397-403.

Weir CR, Hammond KW, Embi PJ, Efthimiadis EN, Thielke SM, Hedeen AN (2011). An exploration of the impact of computerized patient documentation on clinical collaboration. Int. J. Med. Inf. 80(8):e62-e71.

Wu S, Chaudhry B, Wang J, Maglione M, Mojica W, Roth E, Shekelle PG (2006). Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. Annals of internal Medicine, 144(10):742-752.

Xiao X (2008). Privacy Preserving Data Publishing: A Research Summary.

Xiao X (2009). Privacy Preserving Data Publishing: A Research Summary.

Xiao X, Tao Y (2006). Personalized privacy preservation.

Xiao X, Tao Y (2007). M-invariance: Towards privacy preserving re-publication of dynamic datasets.

Yang H, Liu K, Li W (2010). Adaptive requirement-driven architecture for integrated healthcare systems. J. Comput. 5(2):186-193.

Zhong S, Yang Z, Wright RN (2005). Privacy-enhancing k-anonymization of customer data.