ONTOLOGICAL META-ANALYSIS AND SYNTHESIS OF HIPAA

Arkalgud Ramaprasad, College of Business Administration, University of Illinois at Chicago, Chicago, IL, USA, prasad@uic.edu
Than Syn, School of Business Administration, University of Miami, Coral Gables, FL, USA, thant@miami.edu
Khin Than Win, School of Information Systems and Technology, University of Wollongong, Wollongong, New South Wales, Australia, win@uow.edu.au

Abstract

We present ontological meta-analysis and synthesis of HIPAA (Health Insurance Portability and Accountability Act) as a method for reviewing, mapping, and visualizing the research literature in the domain cumulatively, logically, systematically, and systemically. The method will highlight the domain’s bright spots which are heavily emphasized, the light spots which are lightly emphasized, the blind spots which have been overlooked, and the blank spots which may never be emphasized. It will highlight the biases and asymmetries in the domain’s research; the research can then be realigned to make it stronger and more effective. We present an ontology for HIPAA, map the literature onto the ontology, and highlight its bright, light, and blank/blind spots in an ontological map. We conclude with a discussion of how such a map can be used to realign HIPAA research and practice.

Keywords: Ontological Meta-Analysis, Ontological Synthesis, HIPAA.
1 INTRODUCTION

Health information needs to be shared between many health professionals, allied health workers, and associated individuals to enable them to carry out the required tasks in the complex journey of a patient through healthcare diagnosis, treatment, ongoing care, and preventive measures. Health information is also needed for public interest, research, education and other purposes. Defining the correct balance between availability and security requirements of information is a critical goal in a complex environment such as healthcare as excessive procedures to access information will impede the workflow of healthcare providers (Cheow & Win 2009; Win 2005). Maintaining maximum privacy and gaining maximum accessibility at the same time is an important aspect in healthcare industry (Win 2005). Legislations have been developed to protect health information privacy in different countries to ensure health information security, patient privacy, and maintaining accessibility. Different countries have different types of legislation; the Health Insurance Portability and Accountability Act (HIPAA) in the United States, the Personal Information Protection and Electronic Documents Act (PIPEDA) in Canada, Health Record Information Privacy Act (HRIPA) in New South Wales, Australia, are some of the examples (Win & Fulcher 2007). HIPAA was enacted in August 1996 and the Standards for Privacy of Individually Identifiable Health Information (known as the Privacy Rule) became effective in April 2003 in the United States.

Healthcare organizations and relevant entities have adopted business processes to become HIPAA compliant. Both criticisms (Cohen 2008; Frishauf 2005) and adoptions (Cooper & Collmann 2004; Wang et al. 2004) of HIPAA could be seen in literature. Impacts of HIPAA have been studied in different studies (Friedman 2006; Iossifova & Meyer-Goldstein 2012; Murray et al. 2011). Some studies reflect back the time before the HIPAA became effective (Iossifova & Meyer-Goldstein 2012). Since it has been 10 years since HIPAA became effective, the time is ripe to analyse evolution of HIPAA in healthcare industry. Meta-analysis of HIPAA literature will direct the trajectory of the evolution and this will highlight areas that need to be addressed.

The challenge of reviewing and presenting diverse, contradictory, and heterogeneous research literature on HIPAA is large, complex, and ill-structured. Without a clear visualization of the domain one risks replaying the proverbial story of the five blind men each of who imagined an elephant differently after touching its different parts. A sighted wise man helps them map these ‘parts’ and visualize the whole elephant. Similarly, if there is no map or visualization of the research domain the researchers may continue to grope in the dark making the whole less than the sum of the parts. That is the challenge of mapping the research literature in a domain and visualizing it – the metaphorical domain ‘elephant’ is neither fully known nor completely visible. It has to be made known and visible so that its parts can be mapped, the gaps can be seen, and the whole can be made strong, effective, and greater than the sum of the parts. An ontology can help do so (Ramaprasad & Papagari 2009).

Ontology is the study of being in contrast to epistemology which is the study of knowing. Its focus is on objects, their categories, and the relationships between them – the nouns and verbs of the domain. Ontologies represent the conceptualization of a domain (Gruber 2008); they organize the terminologies and taxonomies of a domain. An ontology is an “explicit specification of a conceptualization.” (Gruber 1995, p. 908) It is used to systematize the description of a complex system (Cimino 2006). “Our acceptance of an ontology is… similar in principle to our acceptance of a scientific theory, say a system of physics; we adopt, at least insofar as we are reasonable, the simplest conceptual scheme into which the disordered fragments of raw experience can be fitted and arranged.” (Quine 1961, p. 16) We argue that an ontology is a simple but powerful tool for meta-analysis and synthesis of any research domain, including IS domains such as HIPAA. Cumulative research is important and meta-analysis is an important method to synthesize it. However, meta-analysis is sometimes conducted in a very narrow sense to answer a specific question (for example: Do students learn more when class sizes are small?) or verify a specific aspect of a domain (Hunter & Schmidt 1996). Our method is a holistic approach to survey the landscape and assess the progress of a domain.
(for example: Are we moving in the right direction?) (Noar & Zimmerman 2005). We use meta-analysis in its generic sense that it is the analysis of analysis – that is our analysis is meta or abstraction to that presented in the literature. We distinguish our approach as semantic meta-analysis in contrast to the traditional and more familiar statistical meta-analysis (W. R. King & He 2005) – meta-analysis need not only be statistical. Semantic meta-analysis using an ontology is a new technique we have been developing and presented here in the context of HIPAA. However, it is anchored in the use of ontologies in computer science, medicine, text mining, and linguistics for the analysis of text data.

Computer scientists represent ontologies formally using triples of subject, predicate, and object. We draw upon the concept of ontologies to develop an ontology to envision HIPAA, an important IS domain in healthcare. We will not use the formal triples of computer science but deconstruct the problem into its basic dimensions and corresponding taxonomies incorporating the terminology of the domain. Our ontology is less formal than computer scientists’, more parsimonious than medical terminologists’, and more pragmatic than philosophers’. It is practical, not meta-physical.

We note that we present an ontology of HIPAA not the ontology, recognizing that there can be many equally valid ontologies for the same domain. Each ontology is a lens to study the domain; each lens can offer a different insight about the domain. There may not be a universal ontology for the domain. We can judge the validity of an ontology based on the questions generated from the earlier statements. How explicit is the conceptualization? How specific is the conceptualization? How systematic is the description? And, how systemic is the description? Thus, in the context of the ontology for HIPAA one may pose the questions: Are the dimensions basic to HIPAA? Are the taxonomies basic to the dimensions? And, are the concatenations basic to the domain? We will address these questions using the constructs of face validity, content validity, systemic validity, and external validity commonly used in social science research (Brennan et al. 2011; Horn & Lee 1989).

Ontological meta-analysis and synthesis of HIPAA will provide a method and tools for continuously envisioning the ‘big picture’. The picture can be viewed and analyzed interactively from different points of view and at different levels of granularity. The underlying ontology is itself scalable and extensible and can accommodate future developments in the field. Thus the roadmap for the domain can evolve in synchrony with the emergent science, practice, and needs through continuous feedback (Ramaprasad 1979, 1983) and learning.

In the following we will first present the derivation of an ontology for health information privacy as defined in HIPAA and discuss its validation. Next we present the method of mapping the research on HIPAA onto the ontology. Third, we present the analysis and results in a number of visualizations and associated descriptions. Fourth, we discuss the bright, light, and blank/blind spots in HIPAA research. Fifth, we present our interpretation of the meta-analysis and synthesis of HIPAA research. Sixth, in conclusion we discuss the application of the method to developing a roadmap for HIPAA (and HIPAA-like policies in other countries) research and practice. The above discussion draws heavily upon the work of Ramaprasad (2012) and Ramaprasad and Syn (2013).

2 ONTOLOGY OF HIPAA

The formulation of the HIPAA ontology shown in Figure 1 is based on Ramaprasad and Mitroff’s framework (Ramaprasad 1987; Ramaprasad & Mitroff 1984) for formulating ill-structured problems; it is in turn based on the model proposed by Piaget (1974) for understanding causality. The formulation was manual and not automated. While automated ontology extraction tools such as OWL (OWL 2 Web Ontology Language 2012) are available, they cannot yet formulate an ontology which is (a) parsimonious as the one shown, and (b) organized such that the components can be concatenated as natural language sentences. These tools are designed for standardizing terminologies, as for example in Medicine, but not to extract semantically complete sentences. We used an Excel spreadsheet to organize and present the ontology.
The five dimensions of the ontology are represented by a column each in Figure 1. Each dimension is articulated as a taxonomy of categories constituting it. We will discuss each dimension from right to left. A glossary of dimensions and categories is given in Figure 1. In the following the labels for the dimensions and categories are capitalized. The same words used as part of the discourse are not.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Security</th>
<th>Entity</th>
<th>Outcome</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review</td>
<td>[of] Safeguards</td>
<td>[by] Providers</td>
<td>Confidentiality</td>
<td>Individual</td>
</tr>
<tr>
<td>Identification</td>
<td>Physical</td>
<td>Plans</td>
<td>Integrity</td>
<td>Treatment</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Administrative</td>
<td>Clearinghouses</td>
<td>Availability</td>
<td>Payment</td>
</tr>
<tr>
<td>Mitigation</td>
<td>Technical</td>
<td></td>
<td></td>
<td>Health Care</td>
</tr>
<tr>
<td>Monitoring</td>
<td>Policies and Procedures</td>
<td></td>
<td></td>
<td>Agree/Object</td>
</tr>
<tr>
<td></td>
<td>Organizational Requirements</td>
<td></td>
<td></td>
<td>Incidental</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Public Interest</td>
</tr>
</tbody>
</table>

**Illustrative Components**
Review of safeguards by providers to ensure confidentiality of electronic PHI use/disclosure for individual use.
Evaluation of policies and procedures by plans to ensure integrity of electronic PHI use/disclosure for payment use.
Monitoring of organizational requirements by clearinghouses to ensure availability of electronic PHI use/disclosure for public interest.

**Total number of components encapsulated in the ontology = 1575 (5\*5\*3\*3\*7)**

**Glossary of Dimensions and Categories of the Ontology**
- **Assessment:** Of the information system (IS) to assure security and privacy of personal health information (PHI).
- **Review:** Of data about the security of IS and privacy of PHI.
- **Identification:** Of the key issues (strengths and weaknesses) of IS for privacy of PHI.
- **Evaluation:** Of the impact of the strengths and weaknesses of IS on privacy of PHI.
- **Mitigation:** Of the negative impacts of IS security on PHI privacy.
- **Monitoring:** Of the IS compliance with PHI privacy requirements.
- **Security:** Control of access to PHI.
- **Safeguards:** Measures to control access to PHI.
- **Physical safeguards:** Physical limits on access to PHI.
- **Administrative safeguards:** Administrative limits on access to PHI.
- **Technical safeguards:** Information technology based limits on access to PHI.
- **Policies and procedures:** Expectations that describe acceptable and unacceptable behaviour.
- **Organizational requirements:** Requirements of itself and its business associates to assure security and privacy.
- **Entity:** Organization responsible for assuring privacy of PHI.
- **Providers:** Of healthcare -- clinics, hospitals, and allied organizations.
- **Plans:** Administrators of healthcare plans, insurance, etc.
- **Clearinghouses:** Processors or facilitators of processing of healthcare data.
- **Privacy:** Control of PHI and its use.
- **Outcome:** Control of PHI use.
- **Confidentiality:** Assuring the authorized disclosure of PHI and preventing unauthorized disclosure.
- **Integrity:** Assuring the trustworthiness of PHI and preventing unauthorized modification.
- **Availability:** Assuring the availability for authorized use and prevention of availability for unauthorized use.
- **Use:** Application of PHI for a purpose.
- **Individual:** For use by the patient/health consumer.
- **Treatment:** For treatment of the patient.
- **Payment:** For payment for services.
- **Healthcare:** For maintenance and improvement of physical and mental health.
- **Agree/Object:** For patient to agree or object to the use of his/her PHI.
- **Incidental:** Secondary uses or disclosures that cannot reasonably be prevented.
- **Public interest:** Use required by law, public health, health oversight activity, abuse, negligence, research.

**Figure 1. Ontology of HIPAA**

The ultimate objective of HIPAA is to manage the use/disclosure of electronic personal health information (PHI) without violating the privacy of the individual. The Use dimension lists the seven types of permissible use of PHI. They are: (a) for use by the Individual however he/she chooses to do
so, (b) for Treatment of the individual, (c) for determination of Payment for services, (d) for Health Care of the individual, (e) for the individual to Agree/Object with the contents of the PHI and its use, (f) for Incidental use, and (g) for Public Interest use. The objective of Privacy is to ensure the Confidentiality, Integrity, and Availability of electronic PHI for such use/disclosure – these constitute the categories of the Outcome dimension. The Outcome and Use dimensions together constitute Privacy. Thus, Privacy has 21 components which can be concatenated from Outcome x Use. They include: (a) Ensure confidentiality of electronic PHI use/disclosure for individual use, (b) Ensure Integrity of electronic PHI use/disclosure for agree/object use, (c) Availability of electronic PHI use/disclosure for health care use, and (d) eighteen others.

The entities which have to ensure Privacy are listed under the Entity dimension. They are the Providers, Plans, and Clearinghouses. These entities perform different functions in the healthcare system – they have different needs for and use of PHI. However, in performing their functions they are obliged to ensure each of the 21 components of privacy of PHI through appropriate Security – the next dimension to the left.

The entities are expected to assure security through Safeguards, Policies and procedures, and Organizational requirements. The Safeguards have to be Physical, Administrative, and Technical – they are shown as subcategories in Figure 1. Thus, five categories (including the three subcategories of Safeguards) constitute Security.

The safeguards have to be constantly assessed. The five aspects of assessment constitute the leftmost Assessment dimension. They are: Review, Identification, Evaluation, Mitigation, and Monitoring. Thus there is continuous Assessment of Security by each Entity to ensure Privacy of PHI.

A natural English sentence can be concatenated using a category from each column together with the words/phrases prefixed/suffixed to the column. Three illustrative components are listed below the ontology in Figure 1. Each sentence represents a component of HIPAA. There are 1575 (5*5*3*3*7) components encapsulated in the ontology. They capture the scope and complexity of HIPAA.

Each component may be instantiated in multiple ways. Thus there may be many ways of ‘evaluation of policies and procedures by plans to ensure integrity of PHI use/disclosure for payment use’. It may be done through comparison with best recommended practices, internal assessment committees, and external audit committees. Consequently, some components may be instantiated frequently, some infrequently, and some not at all. We will call the frequently instantiated components the ‘bright’ spots, the infrequently instantiated ones the ‘light’ spots, and the uninstantiated ones the ‘blind/blank’ spots. In the last case it is not possible to determine a priori whether the component has been overlooked (blind spot), or whether it is infeasible (blank spot). The frequency of instantiation of a component may not necessarily indicate its importance, centrality, criticality, or other priority. A ‘bright’ spot may simply be a product of convenience or a herd effect; a ‘light’ spot may be a product of inexperience or oversight; and a ‘blind/blank’ spot may in fact be what the label suggests.

Thus, the ontology is a parsimonious, complete, closed encapsulation of the core logic of HIPAA. Errors of omission and commission in the ontology can be corrected by extending or reducing the dimensions and the categories constituting the dimensions. It is easy to do so because the ontology is modular. By the same token, the granularity of the ontology can be changed by refining the categories or aggregating them. We will use the ontology as a lens to map the extant literature on HIPAA.

2.1 Validity of the HIPAA Ontology

The validity of the ontology will determine the quality of the visualization of the domain and interpretation of the domain map. A number of authors have addressed the issues related to ontology validity and quality in the past (Burton-Jones et al. 2005; Evermann & Fang 2010; Staab et al. 2004). However, their focus is at a finer level of detail, greater formalism, and machine readability than the HIPAA ontology. We, by contrast, draw upon the traditional constructs of validity and assert the face,
content, semantic, and systemic validity of the ontology, and the external validation of the same by the research in the domain (Brennan et al. 2011; Horn & Lee 1989).

The face validity of the ontology is high. It makes sense ‘on its face’ to the experts and novices alike, especially since it is presented using natural English words. Further, the names of the dimensions and categories are part of the HIPAA literature – they have been drawn from it.

The authors have been diligent to include all the elements of HIPAA through an iterative process, and the ontology appears to encapsulate them satisfactorily. Hence the content validity of the ontology is high. Moreover, errors of omission and commission can be easily corrected as described earlier, thus assure content validity even if there has been an oversight.

The components of HIPAA are expressed in natural English sentences, easily understandable to the novice and the expert. Thus each component is semantically meaningful, irrespective of whether it is instantiated or not. Moreover, the users can derive these components from the ontology and judge their meaningfulness. Thus the semantic validity of the ontology is high.

As discussed earlier, the ontology is a complete closed description of HIPAA. Its dimensions and categories are well-founded in the frameworks articulated by the HIPAA governing agency. They are inclusive and yet parsimonious. Thus, it encapsulates all the possible components of HIPAA. Its systemic validity is high.

Last, but not the least, as we discuss below, the ontology is adequate to map the extant literature on HIPAA. No significant dimensions and categories appear to be missing or extraneous. Thus, the ontology has external validity.

3 MAPPING THE CURRENT RESEARCH ON HIPAA

We collected all articles indexed in PubMed between 2003 (when HIPAA went into effect) and 2013 which belong to the MeSH major topic “HIPAA”. The result was filtered for the availability of abstract. We also excluded announcements and letters. The remaining 529 articles were mapped by the authors onto the ontology based on title and abstract.

About 250 of the 529 articles were mapped by all three authors; the rest were mapped by two. Each author mapped the articles independently. The three/two mappings were collated. The coders reviewed the differences and their own mapping in the second round. The final mapping was based on two (of the three) or both authors agreeing. An article could be coded (a) on all or some of the dimensions, and (b) into a single or multiple components of the framework. All the data were maintained and mapped on Excel spreadsheets.

Of the 529 articles, 195 had only an indirect relationship to HIPAA or were general description of the law/rules; hence they were not coded. Of the remaining 334, 5 were mapped onto all the dimensions of the ontology and 329 were coded onto subsets.

4 RESULTS AND ANALYSIS

The results are presented as an ontological map (top of Figure 2) and a heat map of the complete components. The bars in the ontological map are proportional to the parenthetical numbers and represent the frequency of the respective category in the HIPAA research studied. The heat map of the components is color-coded based on the frequency of occurrence of 23 distinct components completely coded; 238 distinct components only partially coded are not shown in the heat map. In the following we will discuss the ontological map and the histogram in greater detail.
4.1 Ontological Map of HIPAA

We will analyze the ontological map of HIPAA research in terms of the ‘bright’, ‘light’, and ‘blind/blank’ categories and components. The analysis is visual and subjective. There are no predetermined frequency bands for the ‘bright’ and ‘light’ spots – the differences are easy to see. The ‘blind/blank’ spots by definition have zero (or very close to zero) frequency. It is a form of gap analysis which is systemic and systematic, and hence more comprehensive than traditional gap analysis. By highlighting all the gaps, both surpluses (more than desired) and deficits (less than desired), it facilitates a synoptic strategy instead of an incremental strategy to the problem of privacy of health information. It is a simple, yet practical, representation of a complex phenomenon.

The ‘bright’ spot of Privacy in HIPAA research is on Public Interest use; the ‘light’ spots are Individual, Health Care, and Treatment use; and the ‘very light’ almost ‘blind/blank’ spots are on Agree/Object, Payment, and Incidental use. For these uses the brightest focus is on Confidentiality followed by Availability, with almost no focus on Integrity.
Among the entities, the brightest – almost exclusive – focus is on Providers; there is very light focus on Clearinghouses, and even lighter focus on Plans.

The brightest focus of Security is on Policies and Procedures, followed by Safeguards – Technical. There is lighter focus on Organizational Requirements and Safeguards – Administrative; and very light focus on Safeguards – Physical and Organizational Requirements.

There appears to be very little focus on Assessment as a whole. Among its categories, there is somewhat more focus on Review and Evaluation than on others.

4.2 Interpretation and Discussion

The research coverage of HIPAA over the past ten years has been patchy. A very large number of categories in and components of the ontology have received little or no attention. Some of them may be unimportant but it would be hard to argue that all of them are. The whole Assessment dimension, for example, has been weakly researched – but yet it is a central part of HIPAA. One could also argue that all healthcare organizations comply HIPAA through Policies and Procedures and Assessment were done regularly but scantily reported in literature. However, the importance of assessment should not be overlooked.

The results of bright spot on Public Interest use of PHI are parallel to the research internationally (Andersen & Storm 2013; T. King et al. 2012; Susilo & Win 2007). Impact on health research, public health, disease surveillance, data linkages reflecting on privacy legislations have been heavily looked at in different countries (Andersen & Storm 2013; Oderkirk et al. 2013). From the results, the ‘bright’ spot, Public Interest use, and the ‘light’ spots, Individual, Health Care, and Treatment also reflected on studies of safeguards and policies and procedures on protecting data in clinical research (Susilo & Win 2007; Win 2005). Technical safeguards involve deidentification, anonymization, privacy preserving data collection, and cryptographic techniques. One can argue that there are several studies on technical perspectives for healthcare organizations and health data exchanges to fulfill HIPAA requirements and ensure privacy and confidentiality of healthcare consumers – for example, privacy and security enhancing dynamic information collection and monitoring (Canim et al. 2012; Xiong et al. 2013). However, those studies heavily focused on technical perspectives and not indexed under HIPAA are not presented in this study. Nevertheless, all articles indexed under HIPAA in MeSH terms are included to comprehensively address the ontological meta-analysis of HIPAA.

Our results pose the question whether the research on HIPAA has been systematic or opportunistic? If HIPAA has to be effective in protecting the privacy of PHI the research should either cover all the components encapsulated in the ontology, or have a rationale for the extremely selective focus resulting in the wide swaths of ‘light’ and ‘blind/blank’ spots in the ontological map. The ‘bright’ spots may be important, but are they exclusively so? It may be argued that many of them are a consequence of the ease of doing research or a self-reinforcing cycle of research and publications without an assessment of their contribution to the systemic effectiveness of HIPAA.

The maps may change over time as HIPAA and its practice evolve. Our mapping can track and guide the evolution, instead of it being ad hoc, by providing a basis for systemic feedback and change.

5 CONCLUSION AND IMPLICATIONS

Ideally research should find application in practice; practice should provide feedback and pose problems for research; and policy should follow research and practice, and provide feedback for both. The patchy ontological map of HIPAA research and the sparse histograms pose two key questions: Is there only a narrow set of research relevant to practice? Is the practice well settled in most components of HIPAA and hence do not require research? Are the policies based on the research and practice strong? We do not believe that the answer is affirmative to any of the three questions. In this context
we believe that the ontology can help direct the research on, practice of, and policies for HIPAA systemically and systematically instead of fragmentarily and opportunistically.

The ontological maps and histograms provide clear visualizations of the gaps. Some of these gaps and definitely need to be bridged. The researchers, practitioners, and policy makers have to assess the importance of the gaps and change research, practice, and policies to bridge them. This process of feedback and change has to be ongoing for continuous improvement of HIPAA. Ontological maps and analysis such as the one presented in this paper can provide the foundations for visualizing the domain, monitoring the incremental changes, and making it complete and integrated.

A major strength of the method is the synoptic view of the domain it provides based on the population of articles elicited by the search. By the same token, its strength is also critically dependent upon the correct specification of the search terms. Another major strength is its use of natural English to model the problem; again the semantic variability of the natural language could also be a weakness. Last, while the parsimony of the ontology is a significant strength, significant errors of omission for the sake of parsimony will weaken the study.

The method is new and in its early stages of development. In the future we propose to systematically address some of the weaknesses mentioned above. We also propose to benchmark the method with other methods for synthesizing the knowledge in a domain.

References


