8-2011

Comparing Requirements from Multiple Jurisdications

David G. Gordon
Carnegie Mellon University

Travis D. Breaux
Carnegie Mellon University

Published In

Proceedings of the International Workshop on Requirements Engineering and Law (RELAW), 2011, 43-49.

Comparing Requirements from Multiple Jurisdications

David G. Gordon
Engineering and Public Policy
Carnegie Mellon University
dggordon@andrew.cmu.edu

Travis D. Breaux
Institute for Software Research
Carnegie Mellon University
breaux@cs.cmu.edu

Abstract

Increasingly, information systems are becoming distributed and pervasive, enabling organizations to deliver services remotely to individuals and to share and store personal information worldwide. However, system developers face significant challenges in identifying and managing the many laws that govern their services and products. To address this challenge, we investigate a method to codify, analyze, and trace relationships among requirements from different regulations that share a common theme of data breach notification. To measure gaps and overlaps between regulations, we applied previously validated requirements metrics. Our findings include a formalization of the legal landscape using operational constructs for high- and low-watermark practices, which business analysts and system developers can use to reason about compliance trade-offs based on perceived businesses costs and risks. We discovered and validated these constructs using five U.S. state data breach notification laws that govern transactions of financial and health information of state residents.

1. Introduction

Modern information systems have grown increasingly distributed and pervasive due to the Internet and wireless communications. Clients, no longer bound by their physical location, can access their data from nearly anywhere - and similarly, that data and the processes that manipulate it can be seamlessly stored or duplicated across multiple servers. With this freedom, however, software developers must engage with an increasingly large number of industry standards and federal regulations that affect their products and services. This leaves developers to contend with a multi-jurisdictional environment, and necessitates new theory to identify a process for achieving regulatory harmony.

A prime example of this can be found in United States data breach notification laws, with 46 of 50 states having such laws at the conclusion of 2010. These laws govern matters such as required levels of encryption, acceptable breach notification, and security procedures. For any developers, especially those belonging to small businesses, distilling these regulations into actionable requirements traceable across their business practices can be difficult. We believe that the existing approach - paper-based laws and policies - can no longer scale with technological innovation, and that the regulations must be accessible to policy makers, business analysts, and software developers if an honest expectation of compliance can be preserved.

As a solution, we believe that regulators and industry can examine regulations as computational artifacts, dynamically linked across jurisdictions. These artifacts can then be integrated with industry and organizational standards to become more easily comparable and addressable. To this end, we put forth an overview of our efforts to formalize a portion of the legal landscape using a requirements specification language (RSL) [BG11] and apply previously validated metrics [BAB08] to compare regulatory requirements using gap analysis. With the results provided by the RSL and gap analysis, we developed operational constructs for high- and low-watermarks to identify and resolve potential conflicts among multi-jurisdictional requirements, and provide system developers with guidance on how to compare regulations. With potential conflicts made salient, system developers and business analysts can consider trade-offs between the additional costs or risks of achieving or forgoing compliance through guided discussions with legal personnel.

The remainder of the paper is organized as follows: in Section 2, we discuss related work; in Section 3, we briefly describe how we encode and compare requirements; in Section 4, we present our case study design; in Section 5, we discuss our research findings, including prominent examples of our watermark construct and trade-offs; in Section 6, we discuss threats to validity; and in Section 7, we conclude with discussion and future work.

2. Related Work

Requirements engineering occurs in the early stages of modern software engineering, wherein terminology is to be grounded "in the reality of the environment for which a machine is to be built" [Jac95]. We now discuss related work in requirements engineering, artificial intelligence, and law.

Requirements specification languages (RSLs), including requirements modeling languages (RMLs), have a rich history in requirements and software engineering [LM82]. RSLs include informal, natural language descriptions to provide readers with context and elaboration, and formal descriptions, such as mathematical logic, to test assumptions across requirements using logical implications [FKV91]. Goal-oriented languages, such as i* [Yu93] and KAOS [DLF93], and object-oriented notations, such as ADORA [GBJ02], include graphical notations to view relationships between entities, such as actors, actions and objects. Because of computational intractability and undecidability of using highly expressive logics [GMB94], RSLs often formalize only a select class of requirements phenomena, e.g., using various temporal logics, including interval [MBK90], real-time [DFL93] or linear [FLM04] temporal logic, or description logic [BAD08]. Consequently, RSLs and RMLs may struggle with the balance between expressability and readability [FKV91].

Unlike i*, KAOS, and ADORA, the RSL utilized herein is designed for the policy domain by integrating formal expressions of document structure with semi-formal

expressions of rights, permissions and obligations, which are required to express regulatory requirements [BVA06]. The RSL emphasizes readability by requiring limited formalization of: actor roles, constraints on those roles, and Boolean logic to express pre-conditions; definitions and their scope of applicability; and cross-references as typed relations between requirements. Finally, the RSL codifies the document structure to ensure certain legal effects from cross-references are traceable and operational – a shortfall of current practices [LG09, MC05, WBM05].

Studies to formalize laws have long been a topic of interest. Early work in the 1980's to encode laws in first-order logic began with a focus on decision support tools [AS84, SSK86], whereas a recent resurgence in formalization of privacy and security regulations have sought to test new theories as expressions of law [DGJ10, MGL06, MA09]. In software requirements engineering, the emphasis is on requirements specification and analysis to develop tools for managing legal requirements. This work has emphasized methodology for encoding laws as rights, permissions, obligations [BVA06], ownership and delegation [GMM05] and techniques for formalizing the legal effects of cross-references, definitions, and exceptions in a comprehensive legal requirements management strategy [Bre09b]. Recent analysis of external cross-references emanating from the Health Information Portability and Accountability Act (HIPAA) shows the potential for conflicts between laws governing different industries [MA11].

Research to compare natural language has long focused on document-level comparisons. K-means cluster [HW79] and latent semantic indexing [DDF90] have been applied to compare documents by examining term frequencies after cleaning the text by removing term suffixes, called stemming [Por80], punctuation, etc. Similar techniques have since been applied to requirements analysis to create traceability links between regulatory requirements and product requirements [CCG10]. In a recent gap analysis between regulatory and product requirements, we discovered that significant domain knowledge is required to recognize semantic differences between requirements, i.e., subsumption, polysemy or synonymy [BAB08]. While tools such as WordNet [Fel98] are used in NLP to supplement domain knowledge for many problems, our research indicates that comparing requirements remains largely a manual process.

3. Encoding Legal Requirements

In preparation to compare regulatory requirements across jurisdictions, we translate the original regulations into a canonical form using a requirements specification language (RSL) [TG11]. The language presents a repeatable and traceable method for documenting requirements within regulations, and provides a number of output formats that are utilized throughout this paper. The requirements are identified used previously validated phrase heuristics [BVA06] and itemized in a standard template. Referenced requirements use the following template: state abbreviation, referenced number assigned by the parser, and then an abridged description of the requirement including who must do what, and under what conditions. Visual representations of requirements used hereafter were generated in GraphML[1] based on the translated regulations. Nodes (requirements) are colored by whether they are permissions (green), obligations (yellow), prohibitions (red) and exclusions (blue) based on annotations.

Regulations from multiple jurisdictions contain potential conflicts due to differences in the administrative hierarchy and requirements coverage. To measure these gaps, Breaux et al. developed and validated a set of statement and phrase-level metrics that an analyst can apply to rationalize document similarities and differences between requirements [BAB08]. For comparing two requirements A and B, the metrics used in this paper are:

Metric S-E (Equivalent): Requirements A and B are equivalent, with some portions of the requirements describing the same or a similar action.

Metric P-G1 (Generalized Concept): The "phrase in B" describes a more general concept than the "phrase in A."

Metric P-G2 (Missing Constraint): The "phrase in A" is missing from Requirement B.

Metric P-R1 (Refined Concept): The "phrase in B" describes a more refined concept than the "phrase in A."

Metric P-R2 (New Constraint): The "phrase in B" is missing from Requirement A.

Metric P-M (Modality Change): The "phrase in A" has a different modality than the "phrase in B."

The process for applying these metrics to itemized requirements proceeds as follows: (1) identify near-equivalent statement pairs A, B and record a logical assertion S-E(A, B); and (2), comparing phrases between statements A, B and record logical assertions P-G1(A, B, p_A, p_B) or P-G2(A, B, p_A) for some phrase p_A in statement A and some phrase p_B in statement B. To compare requirements, the metrics are applied by separately comparing the requirement clauses and the pre-conditions between two requirements.

4. Research Methodology

We now describe our case study research method [Yin08] used to compare multi-jurisdictional requirements from repeated observations of natural language expressions in regulations. The method includes our selection criteria, the translation process, units of analysis, and analysis procedure.

This paper only presents preliminary results towards our goal of observing regulatory variation across multiple jurisdictions and understanding how this introduces complexity into system requirements. We selected a single theme (data breach notification) to limit effects of dissimilarity while we build a new theory to reconcile differences and conflicts among regulations. In the US, 46 state laws were passed between 2003 and 2009 in this domain, and the resulting variations require businesses to reconcile different legally required practices for customers of different states. We selected the following laws by inviting suggestions from a legal expert with seven years of privacy and security law expertise; additionally, Wisconsin was chosen due to its unique inclusion of biometric data as personal information.

- **AR**: *Personal Information Protection Act*, Arkansas Chapter 14.110; 2005.
- **MA**: *Security Breaches*, Massachusetts Chapter 93H; 2007.

[1] http://graphml.graphdrawing.org/

- **MD**: *Personal Information Protection Act*, Maryland Subtitle 14-35; 2008.
- **NV**: *Security of Personal Information*, Nevada Chapter 603A; 2006.
- **WI**: *Notice of Unauthorized Access to Personal Information*, Wisconsin Chapter 134.98; 2006.

Our translation process was conducted by two investigators (the authors) separately classifying statements as definitions, requirements, or exemptions, and writing an expression in the language to characterize the statement. Definitions were identified by common phrases, such as "x means y", where a term x has the logical definition y. Requirements were identified using phrase heuristics identified by Breaux et al. [BVA06] and extended during this study. We maintained a *caveats list* of translation strategies for unusual cases, and a *proposed changes list* of requirements with examples for new language constructs. Laws were reviewed and updated as new constructs were introduced to ensure consistency.

The units of analysis correspond to the translated requirements, definitions, exemptions, and relations between requirements, in addition to the measures produced by the gap analysis. The RSL acts as a natural filter [BG11], capturing only what it can express, which is a threat to validity discussed in Section 6. After the translation, we analyze the units of analysis to identify propositions that link the units to our findings through pattern-based inferences [Cam66]. These patterns consist of constant features (the types of relations and metrics) and the manner by which these constant features structure variable features in the observable phenomena (the different requirements in the relations and the phrase-level measures). We explain the different patterns in our research findings in Section 5.

In the analysis procedure, we first compared similar definitions and then applied the phrase-level metrics to identify dissimilar sub-types and constraints on those types. Second, we compared the requirements by applying the metrics from Section 3 to the requirements clauses and pre-conditions. For requirements clauses measured using the S-E metric, we applied phrase-level metrics to distinguish the differences in terms of *who* is permitted, required or prohibited to do *what*. Next, we consider the dissimilarity between these two requirements in terms of the relations (`REFINES`, `EXCEPT`, etc.) to other requirements. We now discuss our research findings, including the patterns observed through our analysis.

5. Research Findings

The translation of the five laws by two investigators required an average of 2.86 minutes per statement with the first document requiring an average of 2.75 hours or 4.23 minutes per statement, which includes the time to discover the RSL. Each investigator spent an average total of 9 hours to encode the five laws. We observed the number of definitions did not vary greatly and that the number of exemptions was a matter of writing style; neither definitions nor exemptions are proportional to the number of requirements in this dataset.

Our analysis of statements, relations, and measures acquired from the gap analysis yielded several observations. These observations include patterns of dissimilarity, heuristics for reconciling differences and for discovering a legal landscape, and variations in document writing styles that affected our method.

5.1. Patterns of Dissimilarity

When an organization is subject to multiple regulations governing similar business practices, it is likely that the requirements may overlap to some extent by sharing the same subject, action and/or object. Near identical requirements, identified by the S-E metric (without any observed phrase measures) do not pose a compliance issue. However, when the overlap is partial, then the differences between each requirement must be reconciled in order to achieve full compliance. We now outline various differences between requirements and demonstrate by example how an analyst can reconcile these differences. We refer to differences within two requirements as *intra-dissimilarity*, which are determined by comparing the requirement statements using phrase-level metrics. Differences among requirements are referred to as inter-dissimilarity and are determined by comparing dissimilar `REFINES`, `EXCEPT`, and `PRECEDES` relations to other requirements.

An organization must address and reconcile these types of differences before integrating multi-jurisdictional requirements into their systems, policies, and procedures. Normally, this integration is a difficult procedure due to the lack of traceability. However, the RSL and gap analysis offer an improved method for traceability by enabling an analyst to identify, display, and address these differences, incrementally. Consider Figure 1, which shows requirements MD-7 from Maryland §14.3504(b)(2) and NV-9 from Nevada §603A.220(1):

MD-7: a business that concludes the investigation shall notify the individual…

NV-9: a data collector shall disclose the breach to the resident…

Figure 1. Maryland and Nevada disclosure details (abridged)

MD-7 and NV-9 both obligate the entity to notify the individual of a data breach, but their pre-conditions differ significantly: MD-7 requires that the entity conduct an investigation into the breach, whereas NV-9 does not. If it is unlikely that this investigation would interfere with the notification proposed by Nevada, thus an entity might achieve compliance with both regulations by conducting the investigation regardless of the residency of the data subject.

Regulatory requirements may contain thresholds to limit the scope of an obligation. For example, consider MD-18 from Maryland §14-3504(e) and AR-14 from Arkansas §110.105(e)(3) (Figure 2):

MD-18: a business that demonstrates that the cost of providing notice would exceed $100,000 or that the affected class of individuals to be notified exceeds 175,000 may give notification by substitute notice

AR-14: a person or business that demonstrates that the cost of providing notice would exceed $250,000 or that the affected class of individuals to be notified exceeds 500,000 may provide substitute notice

Figure 2. Maryland and Arkansas substitute notice details (abridged)

Both Maryland and Arkansas provide the option of substitute notice when the standard notification methods would be prohibitively complex or expensive. However, these thresholds differ for each state. Due to these quantitative limits, reconciliation to yield one requirement would require choosing the higher Arkansas threshold, thus losing the insight that Maryland residents could be referred to the less expensive substitute notice at lower levels. In such cases, the optimal decision may be to keep the requirements separate and satisfy each requirement, separately.

In addition to intra-dissimilarity observed in phrase-level measures, inter-dissimilarity appears in the presence or absence of relations (such as `REFINES`, denoted with the solid edges) to other requirements. Figure 14 presents a complex example in which three parallel equivalencies are identified (using the double-bar line) between AR-7 and NV-9, AR-8 and NV-10, and AR-10 and NV-12.

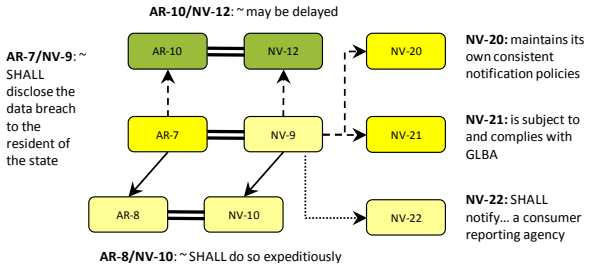

Figure 3. Excerpt from Arkansas and Nevada Comparison (GraphML)

Apart from these three equivalencies, Nevada has additional requirements linked by `EXCEPT` relations (the dashed edge to NV-20, NV-21) and `PRECEDES` relations (the dotted edge to NV-22). The exceptions provide alternative notification mechanisms (comparable internal policies or procedures or compliance with the GLBA). The post-condition NV-22 requires additional notification to consumer reporting agencies to occur after notifying state residents. Because exceptions can halt the discharge of an obligation, the presence of exceptions in one regulation and not another at these equivalencies can cause conflicts. The post-condition, however, is an additional obligation that extends the requirements of the organization, thus they can be treated in the same fashion as `REFINES` relations.

5.2. The Legal Landscape and Positioning

The patterns of dissimilarity illustrate potential conflicts between two regulatory documents as binary comparisons between single requirements. We analyzed the seemingly vast number of comparisons that can be made between all requirements within our dataset, and discovered three heuristics for reconciling differences, which appear in Table 1. Our previous discussion in Section 5.1 presents examples in which these heuristics can be used to resolve potential conflicts or differences between requirements. We believe these heuristics can be applied to potential conflicts across regulatory requirements to discover a legal landscape consisting of choices that system designers must consider in the context of their products and services, business practices, internal policies, preferences, and risk profiles. The borders of the landscape are defined by different standards of care for a finite set of requirements across multiple regulations. A *low watermark* satisfies the minimum requirements by making the fewest decisions in the reconciliation of differences between requirements and occurs when two requirements are precisely equivalent (because neither requirement presumes a higher standard). A *high watermark* is a standard set in which an organization proposes to achieve compliance by the "union" or the "disjoint" separation of differences between requirements. The low watermark standard results from equivalent requirements or the abandonment of relevant details: usually refinements measured by the P-R1 or P-R2 metrics. Alternatively, the high watermark standard seeks to maintain these details in order to achieve or exceed compliance.

Table 1. Heuristics for Reconciling Regulatory Differences

Type	Method
Union	merge expectations (adhering to both if not purely equivalent, or the greater in the case of inclusion)
Disjoint	employ practices that allow adherence to each requirement within its respected jurisdiction
Minimum	determine the floor or lowest common standard

Achieving a high watermark will incur costs beyond those necessary to satisfy the requirements themselves. If dissimilar requirements are reconciled through the use of unions, additional resources will likely be needed given that the covered entities (in this case, additional jurisdictions) will have increased in number. If the two requirements are kept disjoint, we anticipate the need for additional resources (overhead) to maintain separate practices or processes. However, while both of these approaches to dissimilarity resolution result in higher costs, they take on less risk than adhering to the low watermark, minimum standard, which fails to achieve full compliance.

Table 2. Qualities of Watermarks

	High Watermark		Low Watermark	
Decisions	***Union***	***Disjoint***	***Equivalent***	***Minimum***
Compliant	Yes	Yes	Yes	No
Source of Cost	Exceeds Standard	Multiple Standards	Base Costs	Cost of Discovery
Risk	Low	Low	Low	High

5.3. Variation Among Practices

Our process of translation and gap analysis revealed inconsistent styles among the documents. For example, MA §93H placed constraints on what may, must, or must not be done within definitions, as opposed to moving these constraints into rules. Another example includes NV 603A, which lacks an overarching goal to lend direction and context to the document. We now present examples of these consistencies and how they affected our findings.

Common practice within our documents set was to define notice gradually across requirements, leveraging preconditions to add or remove constraints, such as the permission or prohibition for notice to be given through the organization's website. However, MA §93H (1)(a) places these constraints in its definition section. While the definition

describes three types of notice, other regulations have expressed these as permissions refining an original obligation to provide notice. This unusual practice necessitates the comparison of requirements to definitions to thoroughly capture overlaps and conflicts.

We also discovered multiple methods for expressing safe harbors, or regulatory mechanisms that encourage organizations to accept a known outcome or cost in the face of uncertainty. In the RSL, safe harbors can be encoded as exemptions and deference to standards, exclusions (not required to) and "lynchpin" conditions. NV §603A contains many of these practices. Though not shown in the diagram, the regulation contains an exemption (encoded with the `EXEMPT` keyword) for telecommunications providers that excludes them from the requirements shown. Alternatively, a safe harbor is present in this diagram through NV-5, or compliance with PCI-DSS. If an organization is subject to an in compliance with PCI-DSS, the other requirements (NV-6 and NV-7) do not apply to that organization.

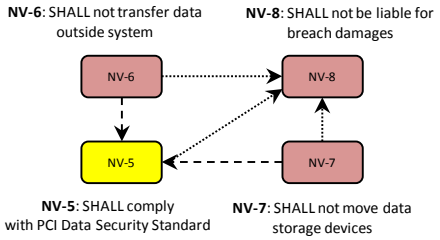

Figure 4. Nevada §603A Safe Harbors (GraphML)

Lastly, perhaps the most obscure type of safe harbor can be found in "lynchpin" conditions. Occurring in definitions and requirements, these conditions, if satisfied, cause the requirement to which they apply - as well as refinements, exceptions, and post-conditions linked to that requirement - to no longer apply. In Figure 5, a number of requirements are traced back to NV-9, specifying how notice be provided, acceptable types of notice, and actions to follow the notification. However, NV-9 has a precondition that restricts the requirement to a breach of unencrypted system data. Thus, if the organization has encrypted their data (based on Nevada's definition of encryption) the requirements no longer apply to the organization.

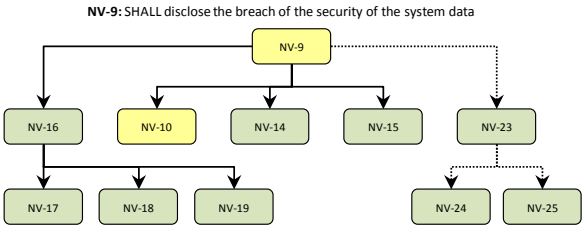

Figure 5. "Lynchpin" Condition in Nevada §603A

6. Threats to Validity

In grounded analysis, multiple analysts derive theoretical constructs from a dataset to describe or explain the data and the constructs are assumed to only generalize to that dataset [GS67]. Recall from Section 4 that we selected regulations that share a theme (data breach notification); thus, our theory may not be externally valid in other regulated domains, such as medical devices or aviation, which may require new language constructs. However, to challenge our assumptions, we validated the schema notation and document model by visually inspecting data breach notification laws in all 46 U.S. states and territories, two U.S. Federal regulations (HIPAA Privacy Rule and Access Standards), the European Union Directive 95/46/EC and a Canadian law (PIPEDA). We found the schema and document model to be sufficiently robust to model these regulatory documents and express their cross-references.

Construct validity is the correctness of operational measures used to collect data, build theory and report findings [Yin08]. To improve construct validity, we maintained a *caveats list* of translation strategies that reflect unusual cases and how the parser should treat such cases, and a *proposed changes list* of requirements with examples for new language constructs. As a new construct was introduced into the language, we reviewed each law to update the translation to reflect the new construct to ensure consistency across the translated datasets. In addition, we developed analytic tools using the parser and a research database to collect all the statistics reported in this paper.

Internal validity is the extent to which measured variables cause observable effects within the data [Yin08]. Our results show that writing styles can positively or negatively impact our methodology, requiring analysts to look beyond the present context to identify dissimilarities between requirements.

Reliability describes the consistency of the theory to describe or explain environmental phenomena over repeated observations [Yin08]. To improve reliability, both investigators (the authors) separately translated the datasets into the RSL and compared their results afterwards to identify alternate modes of expression and language caveats. For the metrics, the investigators compared a subset of their statement equivalencies (S-E measures in the gap analysis) by document pair (e.g. NV-AR, WI-MD, etc.) and determined an initial agreement or "overlap" of over 85%.

7. Discussion and Summary

In this paper, we present the results of comparing five regulatory documents using a requirements specification language (RSL) for codifying legal requirements and qualitative metrics to identifying gaps between requirements. While regulations were not originally written for this type of technical analysis, we believe our analysis can be used to improve the construction of these documents to reach a broader, more participatory audience throughout industry and academia by allowing participation to focus on alternative regulatory structures and their logical implications.

In Section 5, we show how measures of the RSL-encoded requirements can be used to identify patterns of dissimilarity. In addition, we presented heuristics for analysts to use to reconcile potential conflicts between requirements from different jurisdictions. We believe system designers can use the heuristics to select requirements that position their products in better position to achieve full compliance. These selections may be based on costs to design in alternatives based on conflicting requirements, or to choose a common standard that elevates products to a higher standard.

Acknowledgment

This research was supported by the U.S. Department of Homeland Security under Grant Award Number 2006-CS-001-000001, under the auspices of the Institute for Information Infrastructure Protection (I3P) research program.

References

[AS84] L.E. Allen and C.S. Saxon. "Computer aided normalizing and unpacking: Some interesting machine-processable transformations of legal rules." *Computing Power and Legal Reasoning*, pp. 495–572, 1984. West Publishing Company.

[BM07] D. Bourcier, P. Mazzega, "Toward measures of complexity in legal systems." *Int'l Conf. AI & Law*, 2007, pp. 211-215.

[BAB08] T.D. Breaux, A.I. Antón, K. Boucher, M. Dorfman, "Legal requirements, compliance and practice: an industry case study in accessibility." *IEEE 16th Int'l Req'ts Engr. Conf.*, pp. 43-52, 2008.

[BAD08] T.D. Breaux, A.I. Antón, J. Doyle, "Semantic parameterization: a process for modeling domain descriptions."*ACM Trans. Soft. Engr. Method.*, 18(2): 5, 2008.

[BVA06] T.D. Breaux, M.W. Vail, A.I. Antón. "Towards compliance: extracting rights and obligations to align requirements with regulations."*IEEE 14th Int'lReq'tsEngr.Conf.*, 2006, pp. 49-58.

[BG11] T.D. Breaux, D.G. Gordon. "Requirements as Open Systems: Structures, Patterns and Metrics for the Design of Formal Requirements Specifications", Carnegie Mellon University, Pittsburgh, PA, Rep. CMU-ISR-11-100, 2011

[Bre09b] T.D. Breaux, *Legal requirements acquisition for the specification of legally compliance informaiton systems*, North Carolina State Univetsity, Ph.D. thesis, 2009.

[Cam66] D.T. Campbell, "Pattern matching as an essential indistal knowing,"The Psychology of Egon Brunswick.Holt, Rinehart, Winston.pp.81-106, 1966.

[CCG10] J. Cleland-Huang, A. Czauderna, M. Gibiec, J. Emenecker. "A machine learning approach for tracing regulatory codes to product specific requirements." *IEEE/ACM 32nd Int'l Conf. Soft. Engr.*, pp. 155-164, 2010.

[DFL93] A.. Dardenne, S. Fickas, A. van Lamsweerde. "Goal–directed requirements acquisition," *Sci. Comp. Prog.*, 20:3-50, 1993.

[DDF90] S. Deerwester, S.T. Dumais, G.W. Furnas, T.K. Landauer, R. Harshman. "Indexing by latent semantic analysis," *Journal of the American Society for Information Science*, 41(6): 391-407, 1990.

[DGJ10] H. DeYoung, D. Garg, L. Jia, D. Kaynar, A. Datta, "Experiences in the logical specification of the HIPAA and GLBA privacy laws." ACM Workshop on Privacy in Electornic Society, pp. 73-82, 2010.

[Fel98] C. Fellbaum, *WordNet: An electronic lexical database*. MIT Press, 1998.

[FKV91] M.D. Fraser, K. Kumar, V.K. Vaishnavi, "Informal and formal requirements specification languages: bridging the gap." *IEEE Trans. Soft. Engr.*, 17(5):454-466, 1991.

[FLM04] A. Fuxman, L. Liu, J. Mylopoulos, M. Pistore, M. Roveri, P. Traverso. "Specifying and analyzing early requirements in Tropos." *Req'ts Engr. Journal*, 9(2): 132-150, 2004.

[Gar09] B. Garner, *Black's Law Dictionary*, 9th ed, West, 2009.

[GMM05] P. Giorgini, F. Massacci, J. Mylopoulos, N. Zannone. "Modeling security requirements through ownership, permissions and delegation."*IEEE 13th Int'l Req'ts Engr. Conf.*, 2005, pp. 167-176.

[GS67] B. Glaser, A. Strauss. The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine Transaction, 1967.

[GBJ02] M. Glinz, S. Berner, S. Joos. "Object-oriented modeling with ADORA."*Info. Sys.* 27: 425-444, 2002.

[HW79] J.A. Hartigan, M.A. Wong. "A K-means clustering algorithm" *Applied Statistics*, 28(1): 100-8, 1979.

[IEE98] IEEE Std. 1061-1998 – Standard for a Software QualityMetrics Methodology

[Jac95] M. Jackson. "The world and the machine." *17th IEEE Int'l Conf. Soft. Engr.,* pp. 283–292, 1995.

[LG09] M. Lauritsen, T.F. Gordon, "Toward a general theory of document modeling."*Int'l Conf. AI & Law*, 2009, 202-211.

[LM82] A.A. Levene, G.P. Mullery, "An investigation of requirement specification languages: theory and practice."*IEEE Computer*, 15(5):50-59, 1982.

[MA10] A.K. Massey, A.I. Anton, "Triage for legal requirements," NCSU Technical Report #TR-2010-22, October 11, 2010.

[MA09] J. Maxwell, A.I. Anton, "Developing production rule models to aid in acquiring requirements from legal texts." *IEEE 17th Int'l Req'ts Engr. Conf.,* 2009, pp. 101-110.

[MA11] J. Maxwell, A.I. Anton, "Discoverying conflicting software requirements by analyzing legal cross-references," In Submission: *ACM/IEEE Int'l Soft. Engr. Conf.*, 2011.

[MGL06] M.J. May, C.A. Gunter, and I. Lee. Privacy APIs: Access control techniques to analyze and verify legal privacy policies. *IEEE 19th Computer Security Foundations Workshop*, pp. 85–97, 2006.

[MC05] J. Martinek, J. Cybulka, "Dynamics of legal provisions and its representation." *Int'l Conf. AI & Law*, 2005, pp. 20-24.

[MBJ90] J. Mylopoulos, A. Borgida, M. Jarke, M. Koubarakis. "Telos: representing knowledge about information systems," *ACM Trans. on Info. Sys.*, 8(4):325-362, 1990.

[RTA08] S. Romanosky, R. Telang, A. Acquisti. "Do data breach disclosure laws reduce identity theft?" *Workshop on the Economics of Information Security (WEIS)*, June 25-28, 2008.

[Rub11] I. Rubinstein, "Privacy and Regulatory Innovation: Moving Beyond Voluntary Codes." (In Press) *I/S: A Journal of Law and Policy for the Information Society*, April, 2011.

[SSK86] M.J. Sergot, F. Sadri, R.A. Kowalski, F. Kriwaczek, P. Hammond, and H.T. Cory. "The British Nationality Act as a logic program." *Communications of the ACM*, 29(5):370–386,1986.

[Por80] M.F. Porter."An algorithm for suffix stripping."*Program*, 14(3):130−137, 1980.

[WBM05] R. Winkels, A. Boer, E. de Maat, T. van Engers, M. Breebaart, H. Melger. "Constructing a semantic network for legal content," *Int'l Conf. AI & Law*, 2005, pp. 125-132.

[Yin08] R.K. Yin. *Case study research*, 4th ed. In Applied Social Research Methods Series, v.5. Sage Publications, 2008.

[1] [Yu93] E. Yu. "Modeling organizations for information systems requirements engineering." *Int'l Symp. Req'ts Engr.*, 1993, pp. 34-41.