Assessing Regulatory Change through Legal Requirements Coverage Modeling

David G. Gordon
Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, PA
dggordon@cmu.edu

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

***Abstract*—Developing global markets offer companies new opportunities to manufacture and sell information technology (IT) products in ways unforeseen by current laws and regulations. This innovation leads to changing requirements due to changes in product features, laws, or the locality where the product is sold or manufactured. To help developers rationalize these changes, we introduce a preliminary framework and method that can be used by requirements engineers and their legal teams to identify relevant legal requirements and trace changes in requirements coverage. The framework includes a method to translate IT regulations into a legal requirements coverage model used to make coverage assertions about existing or planned IT systems. We evaluated the framework in a case study using three IT laws: California's Confidentiality of Medical Records Act, the U.S. Health Information Portability and Accountability Act (HIPAA) and amendments from the Health Information Technology for Economic and Clinical Health (HITECH) Act, and the India 2011 Information Technology Rules. Further, we demonstrate the framework using three scenarios: new product features are proposed; product-related services are outsourced abroad; and regulations change to address changes in the market.**

Keywords-legal requirements; requirements analysis; legal coverage; regulation modeling; privacy; compliance

I. INTRODUCTION

Organizations that possess or utilize personal information must contend with an increasing number of laws and regulations that include requirements for information privacy and security. These regulations cover a range of functional and non-functional areas, such as the need to encrypt certain data, restrictions on how individuals and third parties may access personal data, and acceptable mechanisms for managing consent prior to collection or use. Costs are not insignificant: the annual cost of data breaches for the healthcare industry alone could be as high as $7 billion in 2013 [20], with implementation and compliance for all U.S. regulations estimated at $1.75 trillion in 2008 [3].

A technical challenge for IT developers is determining when to revisit laws that cover their practices, how to identify relevant legal requirements, and when to examine one's IT systems for compliance with these requirements. Based on our prior experience, we summarize the forces that influence this challenge in the *legal requirements lifecycle* (see Figure 1), which illustrates how legal requirements coverage can change over time. The lifecycle roughly extends Michael Jackson's view that the world consists of both shared and unshared actions with the system [15]; however, our focus is on how changing requirements and laws can affect each other. When developers alter an IT system's requirements and design specifications, we expect to see changes in IT practices described in the world: both changes in how the system is used, and its effects on the world (see Zone 1, Figure 1). Changes in the world, described by facts, lead to new situations that legislators and regulators may not have anticipated, particularly in privacy, where technology and user expectations change frequently. These changes lead to questions about whether existing laws cover new practices and whether this new state of the world is consistent with social, political and economic norms (see Zone 2, Figure 1). If the laws are out of alignment with expectations, lawmakers may propose new laws or amend existing laws to affect *who* and *what practices* the law covers and what prescriptions are written to affect appropriate change. As evidenced by our prior studies [6, 7], these changes to law yield new legal requirements that must be accounted for by engineers, often by modifying their system requirements and designs (see Zone 3, Figure 1).

We more precisely characterize the lifecycle in three sets of knowledge that change with respect to legal requirements coverage: (1) the set R of requirements that consists of non-disjoint subsets L of legal requirements and S of system requirements; (2) the set W of facts about the world, including the context in which the system is developed, deployed, and used; and (3) the set C of conditions that, if true, determine what portion of a law covers the IT system and further entails which legal requirements exist in the subset L. In this paper, we present a method to detect which conditions in the set C entail requirements in the set L.

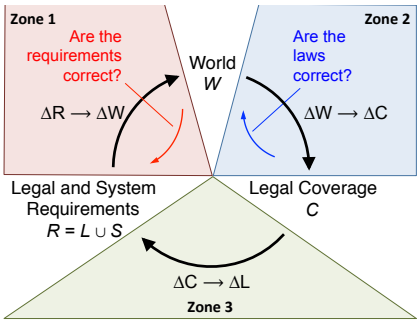

Figure 1. The Legal Requirements Lifecycle: relates changes in legal and system requirements (L, S), the world (W) and legal coverage (C)

To explore the lifecycle, we consider three scenarios in this paper to further motivate and evaluate our approach:

Product Change (ΔR): an organization offers a new product feature or makes significant alterations to features of an existing product.

Moving Abroad (ΔW): an organization moves some or all of their IT processes abroad (e.g., outsourcing), which

produces a change in the facts that describe the processes (e.g., where data is located, how data is processed and by whom, etc.)

Laws Change (ΔC): the laws that an organization is subject to undergo change: a new law is introduced, or an existing law is amended or otherwise altered.

Changes in the legal requirements lifecycle have effects downstream in Figure 1. Software engineers are traditionally concerned with whether $\Delta R \rightarrow \Delta W$, which is to ask, "do these requirements yield the right system," and $\Delta W \rightarrow \Delta R$, "has the world changed in such a way that we must update our requirements to accommodate it?" Alternatively, lawmakers are concerned about other changes: whether $\Delta C \rightarrow \Delta W$, "do these laws correlate with appropriate changes in the world," and $\Delta W \rightarrow \Delta C$, "has the world changed in ways that we need new laws?" These engineering and legal changes occur in parallel and aim to exhibit control over the world, but they do so from intrinsically different perspectives. As such, we believe improving coverage determination (mapping of legal constraints onto software requirements, or $\Delta C \rightarrow \Delta L$), will have important benefits to improving regulatory harmony, while leaving unanswered many questions in how to best translate legal requirements into software requirements and specifications, i.e., $\Delta L \rightarrow \Delta S$.

What do we mean by coverage? In healthcare IT, companies are concerned as to whether their practices are "covered by HIPAA." This broad statement is understood to mean that the HIPAA contains conditions that impose some legal requirements onto the entity. These conditions can vary significantly between laws. In our prior work [13], we observed notable variation among 46 U.S. state data breach notification laws in their *coverage* of organizations, the types of personal information that are *covered*, as well as the criteria under which a breach has occurred or notification must be sent (i.e., the *covered* events). We believe that each scenario described above produces changes in how an organization is covered, as a change in any part of the legal requirements lifecycle can eventually produce a change in *C*.

The remainder of this paper is organized as follows: in Section II, we define relevant terminology; in Section III, we present our coverage model and show how a requirements analyst applies the model to their system context by running example; in Section IV, we state our assumptions underlying the model; in Section V, we present our case study design used to evaluate the model, with our summary findings presented in Section VI; in Section VI we discuss the model's effects on reasoning in three scenarios; in Section VII we address threats to validity; in Section VIII, we present related work, including the modeling of regulatory documents in requirements and other fields, and we conclude in Section IX with a summary and future work.

II. TERMINOLOGY

The following terms are used throughout the paper:

- *Legal Requirement*: an obligation, prohibition, or permission described in a regulatory document that may apply to an organization [9].
- *Pre-conditions*: the antecedent in a logical expression, which consists of properties of an organization, e.g., "owns personal information," organizational roles, e.g., "is a health care provider," or events, e.g., "after a breach of security," that are used to compute regulatory coverage.
- *Assertion*: a claim made by an analyst on behalf of an organization that satisfies a proposition and can be supported by attestation or other evidence.
- *Coverage Model*: structured representation of a natural language regulation into pre-conditions expressed in first-order logic and corresponding legal requirements that are entailed by satisfying propositions using assertions provided by an analyst.

We define legal coverage to mean that an organization is covered by a legal requirement, if the pre-conditions for that requirement are satisfied by assertions; these assertions are documented by the analyst in conjunction with their legal counsel. Based on Black's Law Dictionary [12], coverage is decided by an authority who applies the law to a set of facts; this is typically a judge. While models can be used by requirements and legal analysts to quickly assess coverage early in the design process, we believe they are not a substitute for legal guidance or judicial opinion.

III. LEGAL REQUIREMENTS COVERAGE MODELING

We now present a method to construct the legal requirements coverage model, which analysts can use to check whether a regulation contains requirements that apply to their system context. The method consists of three steps: the analyst: (1) manually translates a regulation text into the legal requirements specification language (LRSL); (2) generates logical expressions from the LRSL-encoded law; and (3) applies the coverage model by making assertions about these logical expressions, assigning truth values to propositions, e.g. "am I a 'body corporate'?" or "do I collect 'sensitive' personal data?" The analyst and their legal counsel make these assertions by collecting evidence about their organization, such as providing relevant attestations or documentation. Using these assertions, the analyst computes the set of requirements that apply to their software system. In this paper, steps 2-3 are novel contributions based on the LRSL, which is prior work [8].

We now describe the coverage model construction method beginning with a brief review of the LRSL using a running example from the India Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules (ITR).

A. Translating regulations into the LRSL

In step 1, the analyst translates the regulatory text into the LRSL for later processing into the coverage model in step 2. The LRSL is based on multiple studies to analyze the syntax and semantics of legal documents [9, 6, 4] and has since been used to study multi-jurisdictional requirements [13]. Consider the following excerpts from the ITR, §§5 and 6 in

Figure 2, immediately followed by the corresponding LRSL encoding:

5. *(1) Body Corporate or person on its behalf... collects sensitive personal data or information...*
   (3) ***When*** *collecting information directly from the person concerned, body corporate or any person on its behalf...* ***shall ensure*** *that the person concerned is having knowledge of... the purpose for which the information is being collected*
6. *(1) Disclosure of sensitive personal data or information by body corporate to any third party* ***shall require*** *prior permission from the provider of such information…* ***unless*** *the disclosure is necessary for compliance of a legal obligation.* ***Excerpts from India's ITR §§5 and 6***

```
1    SECTION ITR.2 //Definitions
2    PAR (c)
3    body corporate     //ITR-A1
4       < company       //ITR-A1₁
5          | firm       //ITR-A1₂
6          | sole proprietorship          //ITR-A1₃
7          | association of individuals   //ITR-A1₄
8             & engaged in commercial or professional
                activities //ITR-A1₅
9    SECTION ITR.5 //Collection of information
10   PAR (1)
11   body corporate
12      | any person on behalf of the body corporate
13      : collects sensitive personal data
14   PAR (3)
15      : collecting information directly from the
          person concerned
16      REFINES (1) #1
17      : shall ensure the person concerned is having
          the knowledge of the purpose for which the
          information is being collected
18      REFINES (3) #1
19   SECTION ITR.6 //Disclosure of information
20   PAR (1)
21   body corporate
22      : may disclose sensitive personal data or
          information to any third party
23      FOLLOWS ITR.5(1) #1
24      : shall require permission from the provider of
          such information...
25      PRECEDES ITR.6(1) #1
26      : it is necessary for compliance of a legal
          obligation
27      PRECEDES (1) #4
28      : must make the disclosure
29      EXCEPT-TO ITR.6(1) #2
```

Figure 2. Example of ITR.5, 6 excerpt encoded in LRSL

The excerpt describes requirements for businesses, called a body corporate, that collect sensitive personal information. The excerpt contains several nested requirements and actions, which an analyst extracts into the LRSL using previously validated heuristics [9]. For example, the modal keyword "shall" (in bold) indicates two obligations. The phrases "when" and "unless" indicate a pre-condition and exception, respectively. In addition to phrase heuristics, we apply previously validated extraction patterns to balance rights and obligations and re-topicalize statements from different actor viewpoints [9].

In Figure 2, the analyst assigns an interpretation to the text by first encoding the section and paragraph references (lines 1, 2, 9, 10, 14, 19, 20) into the LRSL and itemizes each requirement and pre-condition in the same order for which they were discovered: actor roles were separated into logical disjunctions (lines 11-12, where "|" indicates logical-or) and are refined by definitions (lines 3-8, included for context) which are traced across the encoded text by the LRSL parser [8]. Actions follow the ":" colon; actor roles are logically inherited by subsequent nested actions, if no other roles are described. Actions that begin with modal keywords (shall, may) are requirements (lines 17, 22, 24, 28); other actions are often pre-conditions (lines 13, 15, 26).

Actions are linked together using binary relations. In the LRSL, these relations are encoded using special keywords (see lines 16, 18, 23, 25, 27 and 29). Each keyword appears below a requirement (in the relation's domain), followed by a reference that refers to the requirement(s) in the relation's range. We use the following keywords: `REFINES` indicates that a requirement refines another requirement; `FOLLOWS` indicate that a requirement is a post-condition, with `PRECEDES` indicating a requirement is a pre-condition; and `EXCEPT-TO` indicates a requirement is an exception, with `EXCEPT` as the inverse relation. The LRSL assumes multiple relations are linked by conjunction. For a more complete discussion of the LRSL with example regulatory patterns identified in a case study, see Breaux and Gordon [8].

In certain instances, the analyst must re-topicalize a legal statement to represent multiple stakeholder viewpoints. This occurs when rights granted to one party impart obligations or duties on another party [14]. As the coverage model computes coverage for service providers rather than the service consumers, we must apply re-topicalization to ensure that statements about consumer rights are presented from the provider's viewpoint. The analyst may also identify implied requirements from certain phrases. In Figure 2, for example, we inferred the implied permission on line 22 from the phrase "Disclosure of sensitive personal data or information by body corporate to any third party" in the original excerpt. This implied right is balanced by the obligation on line 24.

Translation of the legal text into the LRSL establishes a significant portion of the groundwork necessary to build the coverage model. Once encoded, the analyst uses the LRSL parser to compute a graph, which is the input to step 2.

B. Generating the Coverage Model

In step 2, the analyst generates propositional logic using the LRSL graph created by parsing the previously LRSL-encoded law. To generate the logic, the analyst walks the graph and maps stakeholder roles and non-modal actions to logical antecedents and maps requirements to consequents in a logical implication. If the antecedent or pre-condition in the logical implication is true, we say the system *is covered by* the implicated requirement. We discuss the assumptions underlying this process in section III.D.

To generate the coverage model, the analyst begins with an LRSL-generated graph G that consists of the vertex set $V(G)$ and edge set $E(G)$. The vertices are separately colored as to whether they are non-modal actions (included in pre-conditions) or modal actions (legal requirements), which we refer to as $NON(G)$ and $REQ(G)$, respectively. The directed edges (v, w) correspond to the LRSL relations between these actions; each edge is separately described by whether they correspond to the asymmetric relation for is-refined-by, has-exceptions or has-pre-conditions (e.g., w is a pre-condition of

v, and so on). Thus, the edge subsets are *REFINED_BY*(G), *EXCEPT*(G) and *FOLLOWS*(G), respectively.

Figure 3 presents the LRSL-generated graph from the LRSL-encoded regulation in Figure 2: white nodes represent non-modal actions, black nodes represent obligations, and gray nodes represent permissions; the three types of relations are depicted by dotted edges for `FOLLOWS`, solid edges for `REFINED-BY`, and dashed edges for `EXCEPT`. Suppose we want to ask, when may I disclose sensitive personal data? To answer this question, two propositions are created for the actor roles in Figure 2, which are not represented in Figure 3: the proposition ITR-A1 maps to "is a body corporate," and ITR-A2 maps to "is any person acting on behalf of body corporate." In addition, ITR-20 has a required pre-condition, which is the body corporate must obtain permission from the information provider shown as ITR-21 in Figure 3; we map this pre-condition to *covered_ITR-21*.

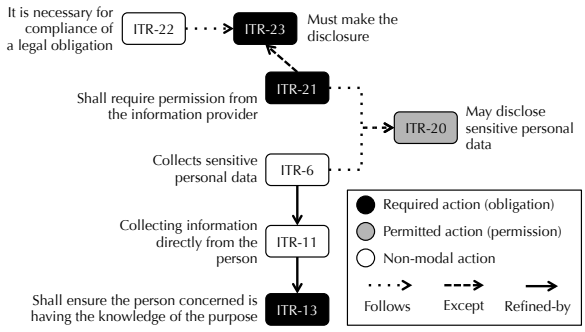

Figure 3. LRSL graph generated from India ITR §5 and 6 in Figure 2. This excerpt focuses on conditions under which sensitive personal data may or must be disclosed to 3rd parties.

Thus, an organization is covered by requirement ITR-20 (a permission), if the following expression evaluates to true:

$$ITR\text{-}A1 \wedge covered\_ITR\text{-}21 \wedge performs\_ITR\text{-}6 \rightarrow covered\_ITR\text{-}20$$

In general, the logic generation process produces the pre-conditions to a requirement through the composition of actor roles, non-modal actions and other requirements that apply before (via `FOLLOWS`) or apply as high-level goals (via `REFINES`). We describe non-modal actions using the proposition *performs_v* for a non-modal vertex $v \in NON(G)$.

The generation process is formalized using a denotational semantics [24], in which we map the LRSL graph notation to first-order, propositional logic, expressed as: $\phi ::= p|(\neg\phi)|(\phi \wedge \phi)|(\phi \vee \phi)|(\phi \rightarrow \phi)$. The denotational semantics consists of a set of valuation functions, below: the double brackets are used to separate the graph syntax from the semantics (quoted fragments of first-order sentences); and the "+" operator means string concatenation. The denotational rules are numbered 1-6 along the left-hand side. The syntactic component $actor[\![v]\!]$ means the logical expression of actor roles for a vertex $v \in V(G)$; thus, $actor[\![ITR\text{-}20]\!] = "ITR\text{-}A1"$ for the example sentence described above. For a vertex $v \in V(G)$, we generate sentences in first-order logic by composing each sentence from substrings, beginning with symbol $expr[\![v]\!]$:

1: $expr[\![v]\!] = lhs[\![v]\!] + " \rightarrow " + rhs[\![v]\!]$
2: $lhs[\![v]\!] = actor[\![v]\!] + edges[\![v]\!]$
3: $edges[\![(v,w) \in FOLLOWS(G) \cup REFINED\_BY(G)]\!] = "\wedge" + rhs[\![w]\!]$
4: $edges[\![(v,w) \in EXCEPT(G)]\!] = "\wedge\ \neg" + rhs[\![w]\!]$
5: $rhs[\![v \in NON(G)]\!] = "performs\_v"$
6: $rhs[\![v \in REQ(G)]\!] = "covered\_v"$

For vertex ITR-20 in the LRSL graph depicted in Figure 3, we apply the denotational rules above as follows (the numbers along the left-hand side indicate which rule was applied to compose the logical expression; rule 0 resolves the $actor[\![v]\!]$ component to the actor role expressed in the LRSL):

1: $expr[\![ITR\text{-}20]\!] = lhs[\![ITR\text{-}20]\!] + " \rightarrow " + rhs[\![ITR\text{-}20]\!]$
2: $expr[\![ITR\text{-}20]\!] = actor[\![ITR\text{-}20]\!] + edges[\![ITR\text{-}20]\!] + " \rightarrow " + rhs[\![ITR\text{-}20]\!]$
0: $expr[\![ITR\text{-}20]\!] = "ITR\text{-}A1" + edges[\![ITR\text{-}20]\!] + " \rightarrow " + rhs[\![ITR\text{-}20]\!]$
3: $expr[\![ITR\text{-}20]\!] = "ITR\text{-}A1 \wedge covered\_ITR\text{-}21 \wedge performs\_ITR\text{-}6 \rightarrow " + rhs[\![ITR\text{-}20]\!]$
6: $expr[\![ITR\text{-}20]\!] = "ITR\text{-}A1 \wedge covered\_ITR\text{-}21 \wedge performs\_ITR\text{-}6 \rightarrow covered\_ITR\text{-}20"$

In addition to actor roles linked to actions, actor roles appear in definitions, where they are defined by sub-types, conditions and examples. Proposition ITR-A1 (body corporate, originally defined Figure 2) is defined in terms of other entities that map to separate propositions, such as a company (ITR-$A1_1$), a firm (ITR-$A1_2$), and so on as follows:

$$(ITR\text{-}A1_1 \vee ITR\text{-}A1_2 \vee ITR\text{-}A1_3 \vee ITR\text{-}A1_4) \wedge ITR\text{-}A1_5 \rightarrow ITR\text{-}A1$$

The complete set of logical expressions generated from the graph in Figure 3 appears in Table I. To infer whether a requirement applies to an organization, the analyst can conduct backward-chaining as proposed by Siena et al. [SIJ+12]. We discuss how to apply the coverage model to a system, which combines forward and backward chaining.

TABLE I. LOGICAL EXPRESSIONS GENERATED FROM LRSL GRAPH IN FIGURE 3 USING DENOTATIONAL SEMANTICS

Ref.	Logical expression
ITR-6	$(ITR\text{-}A1 \vee ITR\text{-}A2) \rightarrow performs\_ITR\text{-}6$
ITR-11	$(ITR\text{-}A1 \vee ITR\text{-}A2) \wedge performs\_ITR\text{-}6 \rightarrow performs\_ITR\text{-}11$
ITR-22	$(ITR\text{-}A1) \wedge covered\_ITR\text{-}21 \rightarrow performs\_ITR\text{-}22$
ITR-13	$(ITR\text{-}A1 \vee ITR\text{-}A2) \wedge performs\_N2 \rightarrow covered\_ITR\text{-}13$
ITR-20	$(ITR\text{-}A1) \wedge covered\_ITR\text{-}21 \wedge performs\_ITR\text{-}6 \rightarrow covered\_ITR\text{-}20$
ITR-21	$(ITR\text{-}A1) \wedge \neg covered\_ITR\text{-}23 \rightarrow covered\_ITR\text{-}21$
ITR-23	$(ITR\text{-}A1) \wedge covered\_ITR\text{-}21 \wedge performs\_22 \rightarrow covered\_ITR\text{-}23$

C. Applying the Coverage Model

We now describe how to apply the coverage model for a new legal context in which coverage has not been previously determined. In our running example, the analyst aims to determine coverage for their organization under India's Information Technology Rules. To do so, the analyst begins with the coverage model *C* from step 2, an empty set *W* of factual assertions about the analyst's system in this legal context, and an empty set *L* of relevant legal requirements that cover their context. In addition, the analyst maintains a table of propositions, their valuation, the assertion(s) that effect that valuation, and any supporting evidence for those assertions, as shown in Table II. This table maintains the analyst's rationale, which can be reviewed later if the system or laws change.

First, our analyst evaluates propositions for the non-modal action ITR-6 in Table I. The antecedent for this modal action includes the proposition ITR-A1 (body corporate), which, as mentioned previously, is a definition

composed of a number of other propositions (e.g. company, firm). As such, definitions are subject to McCarthy or "short circuit" evaluation in which additional propositions need only be evaluated if their truth-value has a bearing on one or more implications. To evaluate propositions, the analyst makes the assertions shown in Table II. The analyst records which proposition is true with respect to their actor role (ITR-A1$_1$, a company), the supporting assertion (W1), and references to any evidence to support the assertion, such as specific forms or documentation the organization maintains (e.g. a patient intake form). Because $ITR\text{-}A1_1 \rightarrow ITR\text{-}A1$, the analyst records the consequent of satisfying ITR-A1$_1$. Based on the implication in Table I, the analyst may perform the non-modal action ITR-6. For each $performs\_$ proposition, analysts make an assertion about whether their organization engages in that action. If true, $performs\_ITR\text{-}6$ indicates the analyst evaluates ITR-11 and ITR-20, such as whether information is directly collected from patients, satisfying $performs\_ITR\text{-}20$.

TABLE II. TRUTH VALUES, ASSERTIONS, AND EVIDENCE FOR BODY CORPORATE AND PERFORMS_ITR-6

Proposition	Val	Assertion	Evidence
$ITR\text{-}A1_1$ (company)	T	**W1**: organization is incorporated under Sections 242 and 245 of the General Corporation Law of the State of Delaware	Delaware Non-Stock Certificate of Incorporation Form
$ITR\text{-}A1$ (body corporate)	T	Implied by $ITR\text{-}A1_1$	N/A
$performs\_ITR\text{-}6$	T	**W2**: organization collects medical information from patient	Incoming Patient Form #8675309A

When the analyst encounters a proposition in the antecedent that has not yet been evaluated, they first consider any propositions that implicate the first proposition (similar to backward chaining). For example, when the analyst proceeds to evaluate the expression for ITR-20, he or she discovers that $covered\_ITR\text{-}21$ must be true, which the analyst has not yet determined. At this point, the analyst first evaluates ITR-21. Once $covered\_ITR\text{-}21$ has been determined, the analyst resumes evaluation of ITR-20, finding evidence to satisfy the antecedent.

The analyst continues making assertions for new propositions as they are encountered, short-circuiting when possible, until all non-modal actions and requirements that appear in antecedents and have not been evaluated (such as ITR-21) are considered. In conclusion, the list of *covered_* propositions corresponds to the set of requirements that cover the organization.

D. Assumptions Underlying the Model

The translation into logic rests on three assumptions $\mathcal{A}1$-$\mathcal{A}3$ that we identified and systematically validated. We now discuss these assumptions with regards to the general depiction of all three relations and their inverses in Figure 4:

$\mathcal{A}1$. Every refinement *y* to a requirement *x* is either

a. A sub-process or task, which means that the act of *y* is temporally contained (begins and ends) within the course of performing *x*, and is necessary to achieve, maintain or avoid *x* (the definition of goal by Dardenne et al. [11]);
b. A quality attribute, which elaborates on either the act of *x* or some object (a noun) within the requirement clause for *x*.

$\mathcal{A}2$. At least one act expressed in the pre-condition *a* must be performed prior to the prescription of y.

$\mathcal{A}3$. For any exception *u* to requirement *y*, if *u*'s pre-conditions are satisfied then *y*'s modality changes to exclusion (e.g., *is required* changes to *is not required*), called a weak exception, and for a strong exception, *y*'s modality changes to prohibition (e.g., *is required* changes to *is prohibited*);

a. Every refinement *z* to requirement *y* inherits this change to exclusion or prohibition; and
b. Every follow-on requirement *b* inherits this change to exclusion or prohibition, respectively.

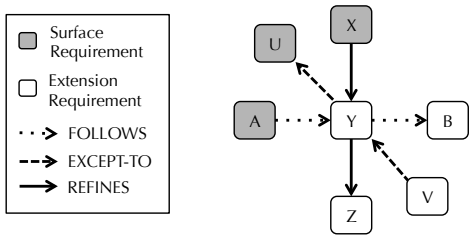

Figure 4. Generic Model of LRSL Relations

We validated these assumptions across a data set of nine U.S. State data breach notification laws, totaling 464 requirements. This was done by automatically generating and manually inspecting traces (*x, y*) and (*a, y*) for assumptions $\mathcal{A}1$ and $\mathcal{A}2$, respectively; for assumption $\mathcal{A}3$, we examined traces (*u, y*), including traces (y, z) for assumption $\mathcal{A}3$ (a) and traces (*a, y*) and (*y, b*) for assumption $\mathcal{A}3$ (b) and, recursively, all corresponding sub-traces from requirements *a, z,* and *b*. Not including recursive traces for assumption $\mathcal{A}3$, this produced a total of 502 traces across all assumptions with the breakdown as follows: $\mathcal{A}1$(a): 194, $\mathcal{A}1$(b): 63, $\mathcal{A}2$: 88, $\mathcal{A}3$(a): 105, $\mathcal{A}3$(b): 52. In this analysis, we encountered obstacles to our approach, such as delay handling, or seemingly ambiguous relationships between requirements that fit multiple relations. For example, one meaning of a goal refinement (how we achieve a goal) is an act performed in preparation for a future event, which may be realized as either a goal refinement or a pre-condition. These discoveries led us to refine our definitions for relations (e.g., to exclude this case from our definition of refinement). Following these revisions and revalidation across the dataset, our analysis indicates that the above assumptions are valid. In future work, we plan to evaluate the extent to which other analysts can apply the method to evaluate these assumptions.

IV. Case Study Design

We now describe our exploratory case study design, including research questions, the data selection, and analysis procedure. Our research questions are as follows:

RQ_1: How do we determine the minimum set of questions for a given law?

RQ_2: How does coverage change when an organization introduces a new product feature, outsources a component of its services abroad, or faces a new or updated law?

We selected the following data sets for our study: India's Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules (2011), California's Confidentiality of Medical Records Act (Civil Code CA §56 et seq., 1981 – 2011), and the United States Security and Privacy Rule from the Health Insurance Portability and Accountability Act or HIPAA (2003 – 2010). Our selection criteria are: (1) we primarily require regulations from multiple jurisdictions that govern both international and provincial, in-nation IT systems; and (2) we specifically chose healthcare IT systems, because healthcare continues to modernize as a significant industry in the U.S. and abroad. Within the U.S., we chose California law because California has the largest expenditure on healthcare by provider of any U.S. state, a measure of productivity and growth that could yield a rapidly innovating market with increased change in regulation (i.e., $\Delta W \rightarrow \Delta C$). For international regulation, we chose India as this nation has grown into a lucrative outsourcing market, including the outsourcing of data processing and medical transcription services. While, these selections are not representative, however, we believe they are significant and illustrate a few of the challenges to solving this problem.

The units of analysis consist of the LRSL-translated requirements and logical expressions generated from the LRSL-encoded graph. We conducted our grounded analysis as follows: the primary author mapped the selected regulations into the LRSL; with logical expressions generated by both authors. To explore the previously mentioned change scenarios, the authors constructed a hypothetical IT healthcare organization based on the Certification Commission on Health Information Technology (CCHIT) ambulatory care requirements, which are used as a standard for certifying healthcare systems in the U.S. The authors applied the coverage models as a requirements analyst responding on behalf of the hypothetical software design team. All examples in subsequent sections were discovered by the primary author and co-reviewed by the second author for errors.

The investigators kept a research notebook to record comments about unusual or notable artifacts in the translation into the LRSL as well as construction of the logical expressions; a list of strategies was recorded to reflect how the investigators handled unusual cases, and upon acceptance of a new strategy, all previous resolutions were reviewed to ensure consistency across the dataset. A legal expert was consulted to review and evaluate findings that arose during the process.

V. Case Study Summary Findings

Herein we describe the timing and scope of effort for our findings. Encoding the regulations into the LRSL consumed 22 hours, with India, California, and the HIPAA accounting for 2.5, 4, and 14 hours, respectively. This initial mapping yielded a total of 394 requirements among the three regulations, attributing 69 to India, 152 to California, and 173 to the HIPAA, reflecting a rate of 2.2, 1.6, and 4.9 minutes per statement. The difference in timing between the HIPAA and other regulations is attributed to the additional time needed to trace definitions within that regulation.

After writing the change scenarios, the authors manually selected one subgraph for each regulation from the fully encoded regulations to demonstrate the effect of changes through the coverage model. These subgraphs contain a total of 98 requirements, summarized in Table III by: the number of non-modal actions (NM), obligations (O), permissions (P) and total per regulation, (T); and the number of binary relations for `REFINES` (R), `EXCEPT` (E) and `FOLLOWS` (F). The three subgraphs contain 21 different actor roles, with the majority (13) coming from the HIPAA. The actor definitions from the HIPAA subsume other actors definitions, which was previously observed as a hierarchy of stakeholders [5].

TABLE III. Requirement and Relation Counts by Regulation

	# Requirements				# Relations		
	NM	**O**	**P**	**T**	**R**	**E**	**F**
CA	4	18	2	24	18	5	3
ITR	6	15	2	23	14	2	10
HIPAA	10	38	3	51	29	5	9

NM: non-modal, O: obligations, P: permissions, T: Total
R: `REFINES`, E: `EXCEPT`, F: `FOLLOWS`

To address RQ_1, the minimum number of questions that an analyst answers increases with the number of stakeholder roles and non-modal actions. For example, the requirements subgraph selected to demonstrate the coverage model in section VI.A for California's regulation contains a total of 7 stakeholder roles; if the analyst claims assertions that falsify all of these roles, no further questions pertaining to non-modals are necessary to establish coverage. This number, which is the lower bound, is based on two assumptions: (1) that the analyst's organization is not covered by the regulation, so responses will be in the negative; and (2) the analyst is able to falsify each stakeholder role with a single assertion, avoiding the propositions that may be introduced by the definition (e.g. for "body corporate" from Figure 2, the analyst claims that his organization is not a body corporate, rather than not a company, not a firm, not a sole proprietorship, etc.) Alternatively, the analyst may satisfy a particular role with a positive assertion and then need to examine non-modals relating to the role. For the upper bound, the analyst must make assertions for all propositions in stakeholder roles and non-modals, avoiding duplicate assertions due to reuse. Table IV presents both bounds for our examples in this paper: notably, HIPAA has a complex stakeholder hierarchy rooted at the role "covered entity".

TABLE IV. MINIMUM NUMBER OF QUESTIONS TO DETERMINE COVERAGE BY REGULATION

	# Questions *(L)*	# Questions *(U)*
CA	7	22
ITR	2	12
HIPAA	2	53

L (lowerbound): *organization is not covered, fewest assertions made by analyst*
U (upperbound): *organization is covered, most assertions made by analyst*

VI. SCENARIO OUTCOMES

We now consider a hypothetical healthcare IT system in three scenarios to evaluate our method: product change, moving abroad and laws change. We grounded these scenarios using five 2011 CCHIT Ambulatory EHR Criteria:

AM 01.01: The system shall create a single patient record for each patient.
AM 02.01: The system shall provide the ability to include demographic information in reports.
FN 04.02: The system shall provide the ability to capture, maintain and display, as discrete data elements, all problems/diagnoses associated with a patient.
AM 26.01: The system shall have the ability to provide electronic communication between prescribers and pharmacies or other intended recipients of the medication order.
AM 39.01: The system shall provide the ability to export (extract) pre-defined set(s) of data out of the system.

The IT system implements these criteria is responsible for managing patient healthcare data acquired from patient admissions, through ongoing care, into patient discharge and billing. For reference, LRSL-encoded actions from India, California, and HIPAA are prefixed with ITR-, CA-, and HP-;. actor propositions include the letter "A" (e.g., CA-A1).

A. Product Change

In the legal requirements lifecycle in Section I, we state that a product requirement's change yields changes in the world, which may cascade through the lifecycle and yield changes in legal coverage. In this scenario, the regulations remain the same, but the analyst must re-examine assertions that led to their prior coverage determination. We consider a clinic, located in California, that wishes to disclose patient information to a third party for research purposes by modifying their system (a product change, **ΔS**) to allow approved third parties to sample anonymous patient data or coded data. The clinic seeks a coverage determination under §§56.05-56.11 of the California Confidentiality of Medical Records Act. Under the Act, health care providers (CA-A1), health care service plans (CA-A2), and contractors (CA-A3) must comply with the unconditional obligation "shall not disclose medical information" in CA-1:

$$(CA\text{-}A1 \vee CA\text{-}A2 \vee CA\text{-}A3) \wedge \neg covered\_CA\text{-}3 \rightarrow covered\_CA\text{-}1$$

Our hypothetical clinic was established as a health care provider (CA-A1), because assertions were made about our clinic's licensure (W3: "is licensed under Section 1200 of the California Health and Safety Code"), and about its systems using requirements AM 01.01, AM 02.01, AM 04.02 as evidence. Therefore, it is covered under CA-1 and $covered\_CA\text{-}1$ is true. However, the regulation describes exceptions to CA-1, such as CA-3 in Figure 5. Many of these exceptions permit information disclosures to certain parties (e.g. a pharmacy) or in light of external events (e.g., receiving a court order). CA-3 allows for general disclosure, allowing the clinic to implement the research requirement. To be covered under this exception, however, an organization must obtain an authorization from the patient for the disclosure ($performs\_CA\text{-}2$), which yields several obligatory refinements governing the authorization (CA-14, CA-19, CA-21, and CA-22, among others).

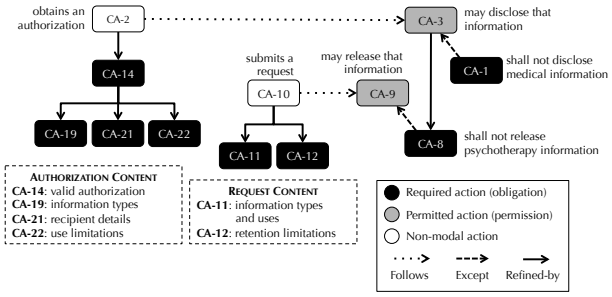

Figure 5. LRSL graph generated from CA §56 Disclosure Requirements. This graph contains authorization requirements necessary to disclose medical information, include exceptions that allow releasing sensitive data types (e.g., psychotherapy information).

Although not unreasonable, the requirement CA-2 to obtain authorization may be burdensome, if each research request entails executing a separate authorization protocol, which must describe the information being collected (CA-19), whom is receiving it (CA-21), and the limitations under which it may be used (CA-22). Further, the clinic that exercises the permitted disclosure is constrained by a refinement (CA-8), which – regardless of the authorization – prohibits disclosure of outpatient psychotherapy information. This refinement may be circumvented through yet another exception CA-9, which requires another type of request for the release of that information in CA-11 and CA-12.

In Figure 6, we present an additional legal requirement: an exception to CA-1 described by the permission CA-5: the disclosure of medical information for research purposes. While this alternative avoids the cumbersome authorization process, it also includes the refinement regarding psychotherapy notes in obligation CA-8. Furthermore, this option is available only to health care providers and health care services, and not to contractors.

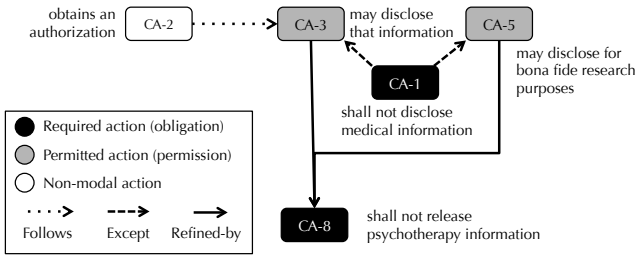

Figure 6. LRSL graph – Coverage Exception for Research in CA §56. Note California's allowance for disclosure of medical information for research purposes, creating an alternative to authorization.

Figure 7 presents the results of applying the coverage model to our product change scenario. The new research-related requirement yields new assertions about the organization's practices (**ΔW**, coming from **ΔS**), which we

traced to changes in legal coverage (**ΔL**). Because the organization was a healthcare provider, they chose the third option that was available in legal requirement CA-5.

ΔS	Usage of medical data for research	
ΔW	Assertions regarding disclosure to third parties for research purposes, evidenced by contracts describing research purposes, constraints on use, etc.	
ΔL	*health care provider* **CA-5**: permission to disclose for research	*contractor* **CA-14 – CA-24**: obligation to obtain authorization for disclosure and refinements detailing that procedure

Figure 7. Impact of Product Change on Coverage Model

B. Moving Abroad

When moving IT services abroad, an organization potentially extends their coverage to include a new jurisdiction. Unlike selling abroad, wherein a product may be prepared to comply with multiple jurisdictions, moving a system component abroad can affect the coverage in both jurisdictions through the removal and addition of legal requirements; particularly requirements that govern the data in that component. Consider the clinic from the previous example, who now wishes to transfer their medical transcription services to India. The immediate effect of moving services abroad is the implication of new requirements from the new jurisdiction. The new requirements can be implied by existing assertions in **W**; i.e., both California and India may regulate the same types of actors and practices. For example, the same assertion that led to the clinic's categorization as a health care provider (W1, W3) are reused to establish it as a "body corporate" under the Information Technology Rules.

In California, a health care provider for which $covered\_CA$-1 was true may exercise permission CA-4 (not shown in Figure 5) to disclose information for "medical data processing" purposes, such as transcription. An obligation CA-7 (also not shown in Figure 5) refines CA-4 and applies to the data processor to prevent the transferred data from being further disclosed.

In Figure 8, we present the relevant portion of India's Technology Rules to our outsourcing scenario. After claiming assertions that classify the organization as a "body corporate" (ITR-A1) that collects sensitive personal data ($performs\_ITR$-6) and collects it directly from the data subject ($performs\_ITR$-11), the clinic is faced with strict requirements regarding the obtaining of patient consent (ITR-7 through ITR-10) as well as ensuring that the data subject is aware of the purpose (ITR-13), recipient (ITR-14), and address of the recipient (ITR-15), in order to reflect the location where the information will be retained. Additionally, the ITR requires the clinic to allow the information provider to withdraw their consent "at any time, while availing services or otherwise".

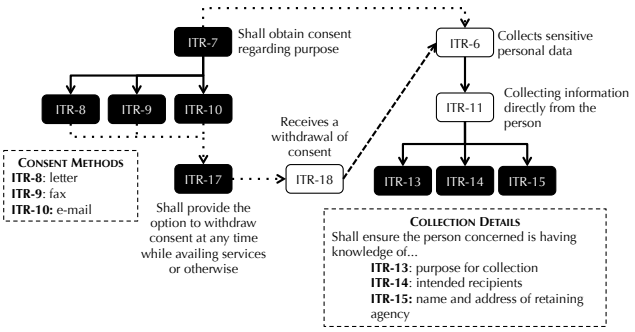

Figure 8. LRSL Graph – Consent Requirements in ITR: India's newest data protection regulation has strict requirements regarding the conferring and withdrawal of consent by the data subject.

Requirements ITR-13 through ITR-15 are similar to and may exceed the authorization requirement CA-2 under Californian law. This is a high water mark effect [13], as the legal requirements imposed through the coverage model under one jurisdiction (India) can influence the operations carried out under another jurisdiction (California). This is shown in the results of applying our coverage model to the moving abroad scenario in Figure 9.

ΔS	Identification of requirements for medical data transcription service
ΔW	Assertions to perform medical information processing in India
ΔL	*body corporate* **ITR-7 – ITR-10**: obtaining data subject consent **ITR-13 – ITR-15**: data subject awareness of purpose and recipient of information **ITR-17**: providing option to withdraw consent

Figure 9. Impact of Moving Abroad on the Coverage Model

C. Regulatory Change

Within a single jurisdiction, regulations may undergo a variety of changes, such as being amended, preempted, overruled – and these actions affect the conditions and requirements that the regulations contain. We now describe a scenario in which the hypothetical clinic adapts to changes introduced by the Health Information Technology for Economic and Clinical Health (HITECH) Act to the HIPAA regulations. Under the old HIPAA, the clinic was classified as a health care provider (HP-A3) and thusly a covered entity (HP-A2) given the definitions for these terms. Among the new requirements that cover the clinic are breach notification requirements (HP-32, HP-35 – HP-41), which are included in the coverage model if the clinic discovers or is made aware of a breach of information ($performs\_HP$-31) and refinements describing the discovery process (HP-24, HP-33, HP-34). Thus, the change in law produces new legal conditions (**ΔC**) and, given the same assertions (**W**) from the prior scenarios, the change implicates new requirements (**ΔL**, as notification requirements) for the clinic.

In addition, covered entities are permitted to disclose protected information to business associates (*covered_HP-43*) or to allow them to create or receive protected information on their behalf ($performs\_HP$-42), only if the covered entity assumes the duty to establish a written contract with business

associates documenting that relationship ($performs\_HP\text{-}44$), and the contract obligates business associates to comply with requirements of the HIPAA Security Rule ($covered\_HP\text{-}51$).

This legal change impacts the clinic by requiring them to revise their business associate agreements, and it also affects legal coverage for the medical transcription service located in India. The medical transcription service who is a business associate (CA-A1) is now effectively covered under requirements in the HIPAA Security Rule, such as:

- **Risk Analysis** (*Required*). Conduct an accurate and thorough assessment of the potential risks and vulnerabilities to the confidentiality, integrity, and availability of electronic protected health information held by the covered entity.
- **Automatic Logoff** (*Addressable*). Implement electronic procedures that terminate an electronic session after a predetermined time of inactivity.
- **Audit Controls** (*Addressable*). [Entity] must implement hardware, software, and/or procedural mechanisms that record and examine activity in information systems that contain or use electronic protected health information.
- **Mechanism to authenticate electronic protected health information** (*Addressable*). [Entity] must implement a mechanism to authenticate electronic protected health information and to corroborate that electronic protected health information has not been altered or destroyed in an unauthorized manner.

As per HP-51, the medical transcription service is not covered by the HIPAA directly, but through a business associate agreement. Unlike the clinic's coverage under India's ITR, wherein requirements were imposed directly upon the organization (after satisfying the stakeholder role "body corporate"), usage of a business associate agreement demonstrates another means of coverage.

VII. Threats to Validity

We now address threats to validity.

Internal validity is the extent to which observed causal relationships exist within the data and, particularly, whether the investigator's inferences about the data are valid [25]. We acknowledge that not all professional analysts will be able to easily perform certain tasks, such as identifying non-modal actions and translating these actions into first-order propositional logic. As such, aim to support our approach with a tool that helps analysts identify these and other constructs, as well as perform the logical transformation.

External validity is the extent to which the framework generalizes [25]. In this study, we investigate portions of three regulations that govern a range of existing business practices. However, this sample is small, and the focus, health IT systems, is narrow. In practice, a single organization's business may span hundreds of jurisdictions, with a single jurisdiction potentially containing hundreds of relevant laws. Although the volume of jurisdictions and regulations in practice represents significant scalability challenges, we believe it also establishes the necessity for a repeatable method to determine and trace coverage changes.

Construct validity reflects whether the construct we propose to measure is indeed what we measured [25]. Before developing our approach, we interviewed a legal expert to understand how they determine coverage under IT laws in practice. The legal expert has experience working with multi-national IT clients in addition to experience with U.S. information privacy legislation. The expert stated that firms customarily rely on in-house experts who have significant experience with particular regulations, and issues of coverage related to those regulations are directed specifically to those individuals. For regulations outside the scope of the firm's knowledge (such as an organization looking to establish operations in an overseas location), the firm performs an analysis of the organization's requirements and conditions and then solicits feedback from outside experts. Based on this interview, the legal construct of coverage is tacit knowledge, but includes considerations such as those we used to model our approach. In future work, we plan to present our coverage model to legal experts to assess whether the coverage determinations are consistent with a legal experts decision-making process.

VIII. Related Work

Laws have long been the focus of formalization by computer scientists. Early work to translate laws into logic include Biagoli et al. [2] and Sergot et al. [21], who expressed legal statutes as logic programs. Formal languages to express laws for legal reasoning include Allen and Saxon's A-Hohfeld language [1] based on Hohfeld's legal concepts [14], and Stamper's LEGOL modeling language [22]. The aim of this research was to support judicial reasoning, in which a court decides whether the facts of a case favor one party or another. Our framework aims to extend this work by incorporating engineering knowledge into decision-making much earlier when design changes are likely to occur and impact legal coverage.

Recent work has been conducted regarding the adoption, interpretation and similarity of law in requirements engineering. In this field, Breaux et al. developed heuristics to extract goals from regulations as rights, permissions and obligations [9, 5]. Our current approach is supported by the legal requirements specification language (LRSL) to enable repeatability [8]. Regulations are interconnected, and Maxwell et al. describe a taxonomy for understanding the impact of legal cross references on requirements [18]. In comparing multiple laws, Lau et al. conducted a similarity analysis of regulations to identify related regulatory texts [16] and Gordon and Breaux proposed requirements "water marking" as an analysis method to identify high and low standards of care across multiple regulations [13]; this method is based on previously validated gap analysis metrics [6]. With regards to assessing proposed laws, Maxwell et al. studied proposed changes to the HIPAA regulations to identify areas of law that were subject to change [19], which may affect requirements changes in our framework.

In requirements engineering, work has emerged to formalize regulations to improve integration between laws and system requirements. Maxwell et al. describe a method to formalize regulations in Prolog [17], which we adapted to express legal definitions in our coverage model. Siena et al. [23] introduce Nómos 2, a conceptual modeling framework that treats regulations as norms and that was applied to the HIPAA Privacy Rule. Nómos 2 allows an analyst to explore regulatory variability by using forward and backward

chaining. While our focus in this paper is on a repeatable coverage modeling method and a scenario-based evaluation to consider three types of change, we believe the Nómos 2 framework could be used to extend our framework further.

IX. Discussion and Summary

In this paper, we introduce a preliminary framework and method that requirements analysts and their legal teams can use to identify relevant legal requirements and trace changes that affect legal requirements coverage. We evaluated the framework in the context of three hypothetical scenarios that yield changes in legal coverage: product requirements change to add new product features; an organization seeks to outsource part of its IT operations; and regulatory change in response to a changing world. While the framework is still preliminary, we believe it is an important step towards discovering ways to systematically identify and respond to requirements complexity in a world that increasingly relies upon multi-jurisdictional software and systems integration.

The framework indicates that significant challenges remain. The number of regulations and rate of regulatory change worldwide suggests that a technical solution will require automation. In future work, we plan to further evaluate the model generation step and enable semi-automated reasoning. This may include using an approach, such as Nómos 2, or an alternative based on Temporal Logic. That said, we believe our study results provide realistic expectations about potential automation: while tools can be brought to bear on processing semi-formal and formal representations, we find the value of these tools is primarily in helping analysts check their assumptions about the meaning of the legal text. In general, we need an appropriate balance between automated tools to encode and reason about regulations and expert legal and requirements advice to make decisions in the context of stakeholder business practices. In the product change scenario, for example, there are multiple choices to implement the new requirement to perform research on healthcare data: each choice yields trade-offs that are within the engineering domain and yet are constrained by laws. In future work, we plan to evaluate our framework in human subject experiments to improve repeatability and identify limitations of tools and human interpretation.

Applying the coverage model to these three scenarios illuminates the need for better integration between legal requirements and architecture. As our product change and moving abroad scenarios demonstrate, adding and removing services can lead to changes in legal requirements coverage. This insight affects how functionality is integrated into a system, where that functionality is located, and what technical constraints affect how that functionality meets stakeholder needs (e.g., performance, security). In the regulatory change scenario, for example, we observe how an Indian service provider might find their services covered by a U.S. law (the HIPAA). To the extent that such changes are predictable and can be isolated in modules of low-coupled software designs, we believe our framework could be used to inform how to limit conflicting regulations by design. We plan to study this dimension in future work through an industry case study with a cloud service provider.

Acknowledgment

We thank Hanan Hibshi and Ashwini Rao for their helpful feedback, and the support of NSF IGERT Award #0903659 and Hewlett-Packard Labs Award #CW267287.

References

[1] L.E. Allen, C.S. Saxon. "Better language, better thought, better communication: the a-hohfeld language for legal analysis." *5th Int'l Conf. AI & Law*, pp. 219–228, 1995.

[2] C. Biagioli, P. Mariani, and D. Tiscornia. "ESPLEX: A rule and conceptual model for representing statutes." In Proc. *1st International Conference on Artificial Intelligence and Law*, pp. 240–251, 1987.

[3] N.V. Crain, W.M. Crain. "The impact of regulatory costs on small firms," U.S. Small Business Administration, Sep. 2010.

[4] T.D. Breaux, *Legal Requirements Acquisition for the Specification of Legally Compliant Information Systems*. Ph.D. Thesis, North Carolina State University, Apr. 2009.

[5] T. D. Breaux, A. I. Antón. "Analyzing Regulatory Rules for Privacy and Security Requirements." IEEE Transactions on Software Engineering, 34(1):5-20, Jan/Feb 2008.

[6] 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.

[7] T. D. Breaux and D. L. Baumer, "Legally "reasonable‘' security requirements: A 10-year FTC retrospective," *Computers & Security*, 30(4): 16-16, Jun. 2011.

[8] T. D. Breaux and D. G. Gordon, "Regulatory requirements traceability and analysis using semi-formal specifications," In Submission: *Int'l Wk'ing Conf. Req'ts Emgr.: Fnd. Soft. Q*., 2013.

[9] 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.

[10] L. M. Christensen, "The paradox of legal expertise: a study of experts and novices reading the law," Brigham U. Ed. and Law J., n.53, 2008.

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

[12] Garner, A. B., Black's Law Dictionary. West Publishing; 9th ed. 2009.

[13] D. G. Gordon and T. D. Breaux, "Reconciling multi-jurisdictional legal requirements: a case study in requirements water marking," *20th IEEE Int'l Req'ts Engr. Conf.*, 2012, pp. 91-100.

[14] W.N. Hohfeld. Some fundamental legal conceptions as applied in judicial reasoning. *The Yale Law Journal*, 23(1):16–59, 1913.

[15] Jackson, M. , "The World and the Machine," *17th International Conference on Software Engineering (ICSE),* 1995, p. 283.

[16] G. T. Lau, K. H. Law, and G. Wiederhold, "Similarity analysis on government regulations," *9th ACM SIGKDD Int'l Conf. Know. Discovery & Data Mining*, pp. 711-716, 2003.

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

[18] J.C. Maxwell, A. Antón, and P. Swire, "A legal cross-references taxonomy for identifying conflicting software requirements," *Req'ts Engr. J*., 17(2): 99-115, 2012.

[19] J.C. Maxwell, A.I. Antón, P. Swire, "Managing changing compliance requirements by predicting regulatory evolution," *20th IEEE Int'l Req'ts Engr. Conf.*, 2012, pp. 101-110.

[20] Ponemon Institute LLC, "Third Annual Benchmark Study on Patient Privacy and Security," *2012*.

[21] 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, May 1986.

[22] R.K. Stamper. "LEGOL: Modelling legal rules by computer." In Proc. *Advanced Workshop on Computer Science and Law*, pp. 45–71, Swansea, United Kingdom, Sept. 1979.

[23] Siena A., Jureta I., Ingolfo S., et al., "Capturing Variability of Norms", *31st Int'l Conf. Conceptual Modeling*, Florence, Oct. 2012.

[24] R.D. Tennent, "The denotational semantics of programming languages," *Comm. ACM*, 19(8): 437-453, 1976.

[25] R.K. Yin. *Case study research*, 4th ed. In Applied Social Research Methods Series, v.5. Sage Publications, 2009.