AN INFORMATION MODELING METHODOLOGY FOR SUSTAINABILITY STANDARDS

Anantha Narayanan
Paul Witherell
Prabir Sarkar
Jae Hyun Lee
Sudarsan Rachuri
National Institute of Standards and Technology, Gaithersburg, Maryland, USA

ABSTRACT

Standards and regulations are developed and introduced in the market to meet the needs of specific domains. As standards are usually developed by experts within a particular domain, the modeling requirements necessary to represent the information associated with these standards are often not well understood. The lack of clear understanding of information requirements creates an environment where information models can become difficult to produce from standards, and the criteria for complying with these standards may be obscure. The variety of challenges encountered in codifying standards using information models necessitates a carefully devised methodology that takes all areas of the whole enterprise into consideration. This paper presents a methodology for the development of information models to complement and support standards based on the Zachman framework for enterprise architecture. In this paper, we will discuss some of the challenges encountered in modeling information for standards and regulations related to sustainability, and subsequently describe how our approach can be used to address these challenges. We will illustrate our approach by developing an example information model to support RoHS (Directive on the restriction of the use of certain hazardous substances in electrical and electronic equipment). This work could lead to the development of software tools and environments for computer aided standards development. Finally, we discuss the advantages and drawbacks of our methodology.

Keywords: information modeling, environmental regulations, Zachman framework

1. Introduction

In the context of products, a “standard” (or “technical standard”) is a common and repeated use of rules, conditions, guidelines or characteristics for products or related processes and production methods, and related management systems practices [1]. Standards may be voluntary or mandatory. Mandatory standards, also called regulations, are rules imposed by regulatory government bodies that specific parties must conform to. Voluntary standards are often used to support regulations by providing a basis for rulemaking and serving as guidelines for conformity. Voluntary standards can serve as a backbone around which information models and a tool support system can be built, which can be used by businesses to work towards conformity.

The recent increase in environmental regulations (especially those originating in the EU, such as RoHS and REACH) has highlighted the need for the accelerated development of information models and tool support systems, in a standardized, repeatable fashion. When the RoHS regulation [2] was enforced by the European Union in 2006, the industry was caught unprepared. Individual manufacturers who could afford the resources implemented ad hoc solutions to support compliance. There was no coordinated effort to create a process that would ensure compliance. One of the main issues for manufacturers was ensuring that the suppliers were RoHS compliant. This issue was addressed in the printed circuit board industry, where RoHS plays a critical role, by the IPC 1752A standard [3]. IPC 1752A is a voluntary standard for materials declaration that establishes the requirements for exchanging material and substance data between suppliers and their customers for electrical and electronic products. It is intended for business to business transactions, and serves as a way for OEMs to monitor substance concentrations in part that they receive from suppliers. IPC 1752A was created with special provisions to address RoHS requirements. NIST developed an information model for the standard [4], along with a prototype tool called SCRIBA [5]. The tool has helped several US manufacturers to monitor their suppliers for RoHS compliance. This is a very good example of how a regulation can be addressed by the industry with the help of a voluntary standard that is supported by a sound information model and software tool.

Unfortunately, it is rarely the case that regulations and standards are supported by open and robust tools. This is
especially true in the case of sustainability-related standards such as RoHS and REACH. The main reasons for this are: 1) these standards and regulations are relatively new and their information requirements are not well understood; 2) they are primarily designed based on human health and environmental perspectives, and may have overlooked the varying impacts on different aspects of manufacturing; and, 3) the domain experts who formulate these standards and regulations are not trained to address information modeling issues.

In this paper we present a methodology for developing information models for environmental sustainability standards based on the Zachman framework. The paper is organized as follows: in Section 2, we give a background of information modeling methodologies in engineering domains, and their application to standards; in Section 3, we explain the complexity of information requirements for sustainability standards; in Section 4, we describe our approach using RoHS as a case study; and in Section 5, we draw conclusions and explore future work in this area.

2. Background

The science of information modeling has developed over the years, with the creation of well defined methodologies to address the information requirements of various domains. Most of the impact has been in software engineering, with the creation of several automated Computer Aided Software Engineering (CASE) tools. We will review some of the modeling methodologies and modeling languages in this section, and study their application for standards modeling.

2.1 Meta-Object Facility

The Meta-Object Facility (MOF) [6] is an Object Management Group (OMG) standard for model-driven engineering. MOF is typically designed as a four-layered architecture, where each layer provides a “meta-model” for the layer below. A “meta-model” defines the syntactic constructs and the syntactic and semantic rules based on which models may be constructed. The models at the lower levels are instances of the models at the level above. The top most layer of the MOF framework, called the M3 layer, is a meta-meta model – it defines the rules for creating meta-models for the layers below. The meta-meta model is defined by OMG. The next layer, called the M2 layer, contains the meta-models for specific domains. The application of the MOF framework for a domain usually begins with the construction of a meta-model for that domain in the M2 layer. The M1 layer contains the domain models, which are based on the meta-model defined at the M2 layer. Finally, the M0 layer contains the physical instances that are represented by the models in the M1 layer. Figure 1 gives an overview of the MOF approach. The dashed arrows (labeled <<instanceOf>>) show which items in the lower level are instances of which items in the meta-model.

The MOF framework is specifically suited for engineering applications that usually involve software-hardware co-design. These are applications such as embedded systems, where abstractions of physical components and software components are modeled together, usually accompanied with automated code generation. These domains are usually well understood and well defined (at least to the extent of the satisfactory functioning of an electronic component). Building tool support for standards must consider a wider array of interacting domains, and does not render itself to a straightforward application of the MOF methodology. We will borrow the ideas of abstraction and meta-modeling from MOF, when describing our holistic approach for standards information modeling.

2.2 UML Modeling

The Unified Modeling Language [7] was developed by the Object Management Group (OMG) as part of their Model Driven Engineering (MDE) program. UML consists of a set of graphic notations, the most popular among which are the UML class diagram, UML state machine diagram and the Use Case diagram.

The UML class diagram is used to model the high level concepts of a domain, along with their attributes and relationships. The state machine diagram is used to model the behavior of a system as a set of states the system can exist in, along with specified transitions between states, which may be triggered by specific events or actions. Use-case diagrams are used to describe specific use cases of the system in terms of the actors and the system functions involved.

UML modeling is most commonly applied in the software engineering domain. The typical methodology followed is to define the system concepts and their relationships using class diagrams, and describe functional requirements using use-case diagrams. Based on these models, the program code and test suites are developed. Though primarily used in software engineering, UML may also be applied to other domains by creating UML profiles. SysML [9] is a UML profile that is commonly used in the systems engineering domain.

UML has been applied to information modeling for
standards in recent times. However, its application has been limited, mainly restricted to defining data exchange. Identifying the information requirements for a tool support system that will help manufacturers in dealing with sustainability standards will require a more holistic approach, as we will describe in Section 3.

2.3 Business Process Modeling

Business Process Modeling (BPM) is the activity of representing the processes of an enterprise for analysis and improvement. BPM is usually performed by business analysts and managers. Technologies such as block diagrams, flow charts, UML and IDEF (Integration Definition) are commonly used for BPM. In recent times, the Business Process Modeling Notation (BPMN) [8] has gained popularity. BPMN is a graphical notation for specifying business processes. BPMN was developed by the Business Process Management Initiative, and is maintained by the OMG.

BPM provides an intuitive way to model business processes for both technical and business users. But its scope is restricted, and does not include other information such as organizational structures or data models. Thus, BPM satisfies only a portion of the requirements for standards modeling. In Section 4, we will describe the use of BPMN for modeling some aspects of the enterprise when creating information models to support standards.

2.4 Zachman Framework

The Zachman framework [10] is a design framework for enterprise architectures. It is a two dimensional matrix that is used to break down the enterprise design concerns into discrete sub-problems, along six cognitive primitives in one dimension, and six levels of detail in the other dimension. This creates a matrix of thirty six cells, where each cell models a portion of the enterprise. Figure 2 shows the cells of the Zachman framework, and briefly describes the nature of the models in each of the cells. While these cells model discrete portions of the enterprise, they combine together to realize the whole enterprise.

The six columns of the Zachman framework are based on six cognitive primitives, namely What, How, Where, Who, When, and Why. The What column is used to capture enterprise information related to data items, the How column for information related to function, and so on, as shown in Figure 2. The rows of the column are based on abstraction of various levels of detail, with the top rows containing high level abstract models, and the lower levels containing more detailed models. While the cells of the Zachman framework provide a clear decomposition of the enterprise, there are no restrictions on the specific models or notations to be used in each of the cells.

The Zachman framework allows us to take a holistic view of the enterprise, while maintaining tractability by careful separation of concerns. We believe that such a framework is ideal for standards, as it allows us to analytically reason about the various disparate issues in dealing with standards. In Section 4, we will describe the use of the Zachman framework for building information models for sustainability standards, and give recommendations for the types of notations that will be suitable for various cells, in the context of providing support for environmental regulations.

<table>
<thead>
<tr>
<th>DATA</th>
<th>FUNCTION</th>
<th>NETWORK</th>
<th>PEOPLE</th>
<th>TIME</th>
<th>MOTIVATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHAT</td>
<td>How</td>
<td>Where</td>
<td>Who</td>
<td>When</td>
<td>Why</td>
</tr>
<tr>
<td>Objective/Scope (contextual) Role: Planner</td>
<td>List of things important in the business</td>
<td>List of Business Processes</td>
<td>List of Business Locations</td>
<td>List of important Organizations</td>
<td>List of Events</td>
</tr>
<tr>
<td>System Model (logical) Role: Designer</td>
<td>Logical Data Model</td>
<td>System Architecture Model</td>
<td>Distributed Systems Architecture</td>
<td>Human Interface Architecture</td>
<td>Processing Structure</td>
</tr>
<tr>
<td>Technology Model (physical) Role: Builder</td>
<td>Physical Data/Class Model</td>
<td>Technology Design Model</td>
<td>Technology Architecture</td>
<td>Presentation Architecture</td>
<td>Control Structure</td>
</tr>
<tr>
<td>Detailed Representation (out of context) Role: Programmer</td>
<td>Data Definition</td>
<td>Program</td>
<td>Network Architecture</td>
<td>Security Architecture</td>
<td>Timing Definition</td>
</tr>
<tr>
<td>Functioning Enterprise Role: User</td>
<td>Usable Data</td>
<td>Working Function</td>
<td>Usable Network</td>
<td>Functioning Organization</td>
<td>Implemented Schedule</td>
</tr>
</tbody>
</table>

Figure 2: The Zachman framework
3. Information Modeling for Standards

Environmental regulations present an interesting challenge in information modeling. These regulations are created primarily in the scope of environmental science, and our understanding of factors detrimental to the environment and human health. However, they are directed towards industry, products and manufacturing processes. It is important to understand the differences in the information requirements from different perspectives to provide a viable support solution for an environmental regulation. Understanding this bridge means not only satisfying the information requirements for conformity, but also addressing the technical and business needs of the industry and helping the industry to adapt and prepare for future challenges.

Let us consider the RoHS regulation. This regulation restricts the amount of six chemical substances (lead, mercury, cadmium, hexavalent chromium, polybrominated biphenyls, and polybrominated diphenyl ether) in electrical and electronics products. The restriction is specified in terms of permitted concentration by weight of homogeneous material (“homogeneous material” is described below). Superficially, the information requirements for this regulation appear to be basic, if not trivial. We only need the weight of the material, the weight of the chemical substance present in it, and the manufacturer can declare conformity with the regulation. However, this overlooks the myriad complexities of manufacturing processes, management of product families, and supply chain issues.

The European Commission defines homogeneous material as a material of uniform composition that cannot be mechanically disjointed (separated by mechanical actions such as unscrewing, cutting, crushing, grinding, and abrasive processes) into different materials. Restrictions on substance concentrations must be met for each homogeneous material in a component, and not just by weight of the whole component. For instance, the hexavalent chromium finish over a base plate could be considered as a separate homogenous material, and must meet the weight restriction without including the weight of the base plating. An information support system must therefore be able to identify all homogeneous materials in a component separately, and record their weights and substance concentrations individually. The information model must be able to account for various types of products with different components, shared components, similar components, etc. The problem is further compounded by variations in similar regulations. China RoHS, for instance, allows reporting substances at the part level instead of the homogeneous material level for certain small objects. A carefully thought out information modeling approach is necessary to meet such complex requirements.

RoHS presents more challenges to deal with. To reach and maintain conformance, manufacturers must monitor their manufacturing processes and supply chain interactions. The information support system must allow manufacturers to track compliance issues closely with their suppliers. Lead soldering is a major contributor to lead concentrations in electronics products. Tin soldering may be used as an alternative. But a side effect of tin soldering is tin whiskers, the formation of tiny metallic hairs on the soldered surface, which could result in short circuiting and failure of the electronic product. Solder baths could get contaminated if there is no process control in place, and lead to non-conformance with the regulation. Parts are sometimes re-soldered during assembly, and could lead to a non-conformant product even if the parts were in conformance earlier in the process. The enterprise model and the assembly model must have the potential to record and capture these issues.

These issues are not unique to RoHS. Restrictions brought in by environmental regulations can have a drastic impact on manufacturing, and variations in manufacturing and assembly processes can impact regulatory compliance. It is clear that a holistic information support system for such regulations must consider various aspects of the enterprise. We will use the Zachman framework to break down and understand various aspects of the enterprise, and devise information models aimed at satisfying the above requirements. The requirements of most environmental regulations are similar at a high level, and this approach is not restricted to the RoHS regulation. The impact of RoHS in recent times has led to some work on information support targeted towards the RoHS regulation. In [11] and [12], Zhou et. al. propose an information system architecture to evaluate product information in PLM system with the RoHS regulation, with the use of XML technology. In [13], Gong et. al. propose an architecture for RoHS risk management to automate information gathering, system integration, and the risk evaluation. While these works are mainly targeted at solving specific declaration and compliance issues related to RoHS, we feel that it is also important to study the breadth of information requirements related to sustainability, and develop a generic methodology that can be extended to cover a wide range of regulations. In this paper, we use RoHS only as an example to describe the methodology.

4. Information Modeling for RoHS: A Case Study

The Zachman Framework allows us to address a complex problem by breaking it into smaller pieces. The value of this framework lies in the dimensions in which the problem is broken down. The question primitives and different layers of abstraction help to break down the problem in orthogonal dimensions, which are easy to comprehend. It also makes it easy to reconstruct the solution from the different pieces by integrating the rows and the columns of the framework. We will use RoHS as a case study to illustrate some of these principles, and provide a representative set of information models that address various aspects of regulatory compliance for RoHS in the electronics industry. We will also try to create general
models from these examples, which may be applied to a wider range of environmental regulations in manufacturing.

Each cell in the Zachman framework contains a model of a portion of the system. These models address finer and finer details as we go down the rows, until the last row which contains the physical system itself. The topmost rows establish the domain of discourse, and identify the important concepts and their relationships. The middle rows provide the logical and physical models that design the enterprise. The lower rows provide the detailed information models that lead to the implementation of the functioning enterprise. We will now look at examples of information models in the different cells. We will use these examples to illustrate the construction of an information modeling solution by problem breakdown and separation of concerns, while also trying to create generalized models that may be applied to other environmental regulations. We describe the type of information model recommended by Zachman for that cell, and how we have applied it for our needs. We describe the notation we chose for that cell and give an example where applicable. In the interest of space, we will limit ourselves to a selected subset of the cells in the framework.

4.1 Top layers: business interest

The top rows of the framework serve to establish the domain of discourse. They will identify the main goals of the enterprise, the most important concepts in the domain, the high level processes in the domain, and the people and organizations responsible for various activities.

Zachman recommends that the cells in the first row, called the contextual row, are simple lists of words. They serve to identify the scope of the problem. The second row, called the conceptual row, is a basic entity relationship model that provides the enterprise model for business decision making and planning. Let us consider the Why column. The first cell in the Why column simply contains a list of business goals and strategies. This could include items such as “RoHS compliance” and “create profits,” and also more thoughtful items such as “brand image” and “corporate environmental responsibility.” The second row models various business plans based on the business goals, in the form of a high level ends- vs.-means model. For instance, the goals (ends) “create profits” and “RoHS compliance” could be related with activities (means) such as “modifying assembly process” and “improving material choices.” This provides a basis for high level business discussion and decision making, without being lost in the details of specific chemicals or manufacturing processes. We use the What and How columns to give more concrete examples below.

Similar to Why, the first row of the What column models the important concepts in the domain, simply as a list of nouns. To construct this model, we began listing high level concepts relevant to RoHS, which were gleaned from a variety of documents and papers related to RoHS. We then converted the listed items into a generalized list of nouns by elevating the concepts to a uniform level of abstraction. This laid the foundation for the What component of the information framework. Table 1 shows an abridged list of nouns relevant to RoHS. As part of our future work, we plan to use this approach to construct an exhaustive taxonomy that will serve as reference data for the information analysis of other environmental regulations.

Table 1: Generalized nouns for What

<table>
<thead>
<tr>
<th>#</th>
<th>Original nouns from documents</th>
<th>Generalized nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Electrical and electronic equipment</td>
<td>Product</td>
</tr>
<tr>
<td>2</td>
<td>Hazardous waste</td>
<td>Hazardous waste</td>
</tr>
<tr>
<td>3</td>
<td>Hazardous substances</td>
<td>Hazardous substances</td>
</tr>
<tr>
<td>4</td>
<td>Landfill sites</td>
<td>Earth</td>
</tr>
<tr>
<td>5</td>
<td>Environmental contamination</td>
<td>Environmental performance</td>
</tr>
<tr>
<td>6</td>
<td>Equipment category</td>
<td>Product category</td>
</tr>
<tr>
<td>7</td>
<td>Exempted category</td>
<td>RoHS category</td>
</tr>
<tr>
<td>8</td>
<td>Directive 2002/95/EC</td>
<td>Sustainability Regulations</td>
</tr>
<tr>
<td>9</td>
<td>Economic Operators</td>
<td>Economic Operators</td>
</tr>
<tr>
<td>10</td>
<td>Human health</td>
<td>Environmental performance</td>
</tr>
<tr>
<td>11</td>
<td>Penalty of non-compliance</td>
<td>Economic impact</td>
</tr>
<tr>
<td>12</td>
<td>Cost of compliance</td>
<td>Economic impact</td>
</tr>
<tr>
<td>13</td>
<td>High Risk Components</td>
<td>High Risk Components</td>
</tr>
<tr>
<td>14</td>
<td>XRF analysis</td>
<td>Testing method</td>
</tr>
</tbody>
</table>

Figure 3: Conceptual Data Model
This model takes the form of a simple entity relationship diagram. To construct this model, we took the nouns identified in the first row, and determined the relationships between them. We have identified some key relationship types such as “belongs to” and “measures,” from which we can create a meta-model for this cell of the Zachman framework, which can be generically applied across various regulations.

The **How** column is used to model processes, or how tasks are performed. The first two rows contain high level process models covering both technical and administrative tasks. For these rows, we have chosen a notation that combines BPMN and IDEF0. The first row of this column models a list of processes relevant to the business. We chose to model the relevant processes as “process chips,” identifying the high level processes and their associated high level information. We will build a library of process chips that will serve as a reference for other environmental regulations. Figure 4 shows a process chip called “Test.” Using the IDEF0 notation, we denote that the arrows coming from the left are the inputs for this process and the arrows going right are the outputs. The arrows coming from the top denote controls, and the arrows at the bottom denote reference material for this process. The process “Test” in Figure 4 takes items such as “aggregated product data” and “materials declaration” as inputs, and gives items such as “test report” as outputs. The information for these inputs and controls comes from the other columns of the Zachman framework, maintaining the separation of concerns.

The second row of the **How** column provides the business process model for the processes identified in the first row. We use BPMN to give the details of the process chips we have defined. Figure 5 shows the business process model for the process chip Test. This models the typical workflow for a test scenario for environmental regulations. It consists of sub-modules for reviewing test information, executing the test and reporting.

The other columns contain models of abstraction level similar to the ones discussed above. The **Who** column models the list of contacts, and the organizational work flow. The **When** column models the list of events and the master schedule for the enterprise. The **Where** column models the list of relevant geographical locations, and the high level logistics network of the enterprise. The separation of concerns and the high level abstractions provided by this portion of the Zachman framework allows the high level executives and business
decision makers in a company to understand, analyze and plan their response strategy to environmental regulations. Our application of the Zachman framework for environmental regulations will involve generic models and extensive reference data for the kinds of information models suitable for these cells of the framework.

4.2 Middle layers: design interest

The third and fourth rows of the framework contain the “logical” and “physical” models of the enterprise. We will use these rows for the system and technical models for the design of the enterprise. We believe that the models in these rows will contain technical design information at a finer level of detail, but not the complete implementation details.

The third row of the What column contains the logical data model. We have chosen the Core Product Model (CPM) notation [14] for this cell. Figure 6 shows an example model for the RoHS regulation. This model uses the data items defined in the first row of the What column, along with the relationships defined in the second row (semantic model). CPM was specifically developed to model the form, function and behavior of products, and allows the association of functions to form. We believe that this model can be used to capture the high level ways in which an environmental regulation can impact the design of a product.

The fourth row of the What column contains the physical data model. We use UML class diagrams for this cell. Figure 7 shows a class diagram describing the information requirements for materials declaration for the RoHS regulation, and is part of the IPC-1752A standard. We can consider this model as a specialization of the CPM in the third row. It will form the basis for the data definition in the fifth row, which we will describe later.

4.3 Bottom layers: implementation interest

The fifth row of the framework contains the detailed models required for the implementation of the system. The sixth row does not contain any models, but represents the working system itself. These last rows of the framework are required to realize the implementation of the enterprise. The models at this level are usually very specialized for the specific requirements of the enterprise. While we may not have models in this row that will apply at a generic level for all environmental regulations, the process of breakdown and construction using the Zachman framework can still be applied for other cases. The last row of the framework represents the functioning enterprise itself. If the framework has been implemented correctly, the functioning enterprise will satisfy all the specifications that were identified in the top five rows.

Figure 6: CPM for What column

Figure 7: Data model for IPC-1752A [3]
The fifth row of the What column contains the data definition model. For instance, this may be database definitions in SQL, which may be derived from the class diagram model in the fourth row of the What column. The fifth row of the How column may contain artifacts such as instruction manuals for running tests or programmed instructions for specific manufacturing activities, and the fifth row of the Where column may be the detailed specification of the network architecture.

5. Conclusions and Future Work

An information support solution for environmental regulations in manufacturing must address various engineering, logistic and administrative concerns that may not have been foreseen during the drafting of the regulation. The Zachman framework provides a method for breaking down the problem by separating concerns, and developing an information support solution. Generalized models based on internationally approved standards will provide a sound basis for the creation of robust software solutions to support regulatory compliance.

Our plan for future work is to create semi-formal descriptions of the overall methodology of applying the Zachman framework for developing information models to support environmental regulations. This will also involve creating more formal and generic meta-models for the individual cells of the framework, which may be applied to a range of environmental regulations. An extensive set of metamodels and reference data will be created, which will cover a wide range of information related to environmental regulations and manufacturing. While we restrict ourselves to the domain of standards and regulations related to sustainability, the same approach may be used on a variety of other general standards. We anticipate that this will result in software tools similar to Computer Aided Software Engineering (CASE) tools, specifically developed to support standards compliance in manufacturing.

Disclaimer

Certain commercial software products are identified in this paper. These products were used only for demonstration purposes. This use does not imply approval or endorsement by NIST, nor does it imply that these products are necessarily the best for the purpose.

References