Automating Conventional Compliance Audit Processes

Johannes Dimyadi, Robert Amor

HAL Id: hal-01764201

https://hal.inria.fr/hal-01764201

Submitted on 11 Apr 2018

Automating Conventional Compliance Audit Processes

Johannes Dimyadi and Robert Amor

Department of Computer Science, University of Auckland, Auckland, New Zealand
{jdim006, r.amor}@cs.auckland.ac.nz

Abstract. Any product, especially those with safety features or concerns, is normally subject to compliance audit with various standards and legal requirements at different stages throughout its lifecycle. These requirements are typically conveyed in voluminous written natural language texts requiring much expert interpretation. The compliance audit process has conventionally been a manual undertaking, which is known to be laborious, costly, and error-prone. In an era of increased legislation and electronic representations of products, it is prudent that some of these manual processes should be automated. This paper describes the capabilities of an automated compliance audit framework that can be incorporated into the compliance management of a product lifecycle. Apart from the product data model that is subject to audit, essential components of the framework include machine-readable legal knowledge and executable audit process models, support for supplementary human input and interface with simulation tools.

Keywords: automated compliance audit processes, computable legal knowledge, building information modelling

1 Introduction

The lifecycle of a product from inception, through design and manufacture, to service and end-of-life stages, is subject to compliance audit processes one way or another, be it legislative, regulatory, or contractual. For example, major purchasers often impose on their suppliers to be ISO 9001 certified to demonstrate that their operation is in accordance with specified quality management system (QMS) requirements. This means there are clearly defined quality control processes in place that can be used to audit the production and the products for compliance with applicable standards.

In the domain of AEC/FM (Architecture, Engineering, Construction & Facilities Management), any component of a building structure is a product that must be designed to comply with legal requirements governing safety and quality standards before it can be consented for installation or construction. The compliance audit process has conventionally been a manual undertaking, which is inefficient, resource intensive and error-prone. A study in New Zealand has shown that 10% efficiency gain in the domain would bring about a 1% boost in the annual gross domestic product [1], which is currently worth in excess of USD75 million per annum.

There have been numerous research attempts over the last half of the century to automate compliance audit processes in the domain [2, 3], but there are only a handful of successful implementations reported to date [4, 5] and most of them only have limited applications. The main challenge remains with accessing and processing the right information efficiently and effectively. Legal knowledge, in particular, is conveyed in voluminous paper-based documents in natural language text written for human interpretation. Recent advances in the fields of artificial intelligence has made, for example, natural language processing techniques available for machines to extract and interpret legal text [6]. However, this approach is not yet matured sufficiently to be fully exploited in real-world applications. More importantly, however, not all legal text can be processed by machines alone. There are implicit types of legal knowledge that only human experts can understand and interpret in a certain context, especially under extraordinary circumstances. Human experts are equipped with intuition and tacit knowledge, and have the ability to draw on their years of experience to make judgments and conclusions reasonably quickly. Moreover, many types of legal knowledge cannot be predefined easily due to their dependency on dynamic factors, which must be evaluated by dedicated computational or simulation processes.

The conventional compliance audit process is procedural in nature, which lend itself to automation. However, there are still roles in the process that are best played by human experts such as specifying what information to retrieve from which sources and how to process them. Machines excel in executing instructions efficiently and accurately and so should be given such a role to play in the process.

Ultimately, there needs to be a practical compliance audit framework where human experts can specify the audit procedure each time, and the product model and legal knowledge are both treated as independent input components to the process engine.

1.1 Human-guided Automation Process

An essential ingredient of the aforementioned practical framework is maintaining direct human involvement in the audit process by allowing human experts to specify the correct type of information to retrieve and sources where they can be retrieved from, and also how they can be processed. The tasks involved are already part of the conventional manual design procedure, so it should be relatively straightforward to transfer the knowledge across for machine processing. One method is to formally document the procedure in a standard process model that can be used as an input component to the framework, which can then be executed in the computing environment, thereby automating the conventional process.

1.2 Product Data Model (PDM) and Building Information Modelling (BIM)

A PDM (Product Data Model) defines the structure of a product model, which is a source and repository of information on the development of a product [7] and a subset of PLM (Process Lifecycle Model) that is primarily concerned with information exchanged during the design phase of the product's lifecycle. PDM is the subject to be audited for compliance. In the context of this paper, this is taken to be equivalent to the

compliant building design data captured in a building model developed using the open standard BIM (building information modelling) approach.

The open standard method of exchanging general product data is ISO 10303 STEP (Standard for the Exchange of Product model data. In the context of the AEC/FM domain, the open standard specifically for exchanging generic building data is ISO 16739 IFC (Industry Foundation Classes) [8]. For a specific application such as the compliance audit in a sub-domain, a subset of the IFC schema known as the model view definition (MVD) can be used to exchange selected and targeted information [9]. As in the conventional practice, separate MVDs can be used to represent different design packages (or product models) related to different design disciplines such as architectural, structural, fire safety, and so on.

1.3 Human and Machine-Readable Legal Knowledge Representation

A legal knowledge representation is what PDM/BIM must be audited against. As an ingredient of the framework, there needs to be a standard digital version of all legal documents that are human and machine-readable. For interoperability, it is pertinent that an open standard model is used to represent these documents.

Two emerging open standards, i.e. LegalDocML [10] and LegalRuleML [11], being developed by OASIS (Organisation for the Advancement of Structured Information Standards), have the potential of being used for this purpose. An important feature of these standards is that they allow the literal and logical content of any paper-based document to be represented and maintained coherently. Furthermore, they support data exchange in open standard XML (Extensible Markup Language), which further promotes interoperability.

2 Automated Compliance Audit Framework

The automated compliance audit framework (Fig 1) described in this paper is capable of executing instructions embedded in a process model, referred to as CDP (compliant design process), and processing information retrieved from the PDM/BIM (product/building model or MVD), LKM (legal knowledge model), and other supplementary sources (manual human inputs and external simulation processes).

The core of the framework is an audit engine that incorporates a dedicated BPMN-compliant process engine. LKM and PDM/BIM are queried by the process engine like any standard information models for compliance assessment in accordance with the specification and instructions in the CDP.

Two optional input components of the framework are manual human inputs and any specified outputs returned by external computation or simulation processes [12]. These supplementary inputs are often necessary to supply information that may be missing from the PDM/BIM or LKM, or to help determine parameters that cannot be predefined or are dependent on dynamic environmental conditions such as air temperatures.

The output of the framework is a set of reports for each CDP highlighting any violation or compliance items that cannot be automatically determined and that may require further attention or processing.
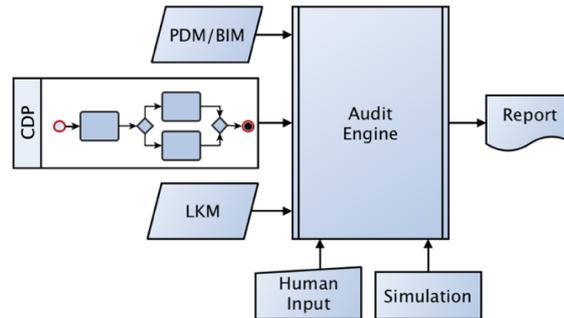

Fig. 1. Automated compliance audit framework (adapted from [13])

3 Executable Compliant Design Processes (CDP)

A building design is subject to multiple compliant design processes relating to different building components or aspects such as architectural, structural, fire safety, electrical and mechanical services, or others. In conventional practice, designers typically follow industry standard procedures, which may or may not be documented, to develop various design solutions that are compliant with applicable standards and regulatory requirements. The procedural design task is amenable to automation as each step of the procedure can be mapped into respective activities in an executable process model.

3.1 CDP as A Component of Legal Knowledge

Legal documents contain provisions that may not all be applicable to every situation. There are multiple compliance paths present in each legal document. Choosing a set of scenarios leads to a particular path of compliance. Selecting a different set may result in an entirely different path. Each compliance path represents a compliant design procedure that can formally be documented as a CDP for computer execution. Each CDP is, therefore, a component of the legal knowledge conveyed by the document.

It is the designer's role to evaluate and decide which compliance path to follow (or which CDP to use) and whether or not the resulting compliant design is deemed satisfactory. Such a decision making process typically requires an intimate knowledge of different compliant design options, considerations of cost-benefits, the acceptable level of risks and safety margins. These are human attributes that are difficult to transfer to machines. On the other hand, machines are far superior than humans in terms of executing repetitive instructions efficiently and accurately. Therefore, it is considered appropriate and natural for human experts to resume an active role of providing direct guidance to the automation process by means of instructions specifying sources of information and what information to retrieve, and how the information is to be processed.

Machines are simply given the role of doing what they do best, which is to execute specified instructions.

3.2 BPMN-compliant CDP

The sequence of steps in a typical compliant design procedure can be represented as a series of activities, events, and sequence flows in a process model such as the open standard Business Process Model and Notation (BPMN) [14], as shown in Fig 2. The current BPMN standard (BPMN 2.0) promotes extensibility and interoperability by supporting data exchange in XML natively. One important type of activity in a BPMN-compliant process model is the script task, which allows embedding of computer scripts that convey user-specified instructions such as where to retrieve which information from and how to use the collected information to perform specific calculations.
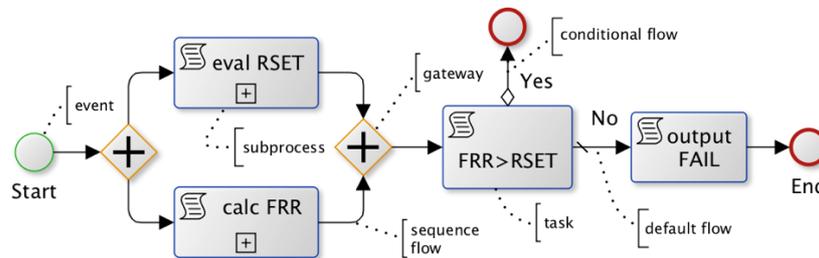

Fig. 2. Exemplar BPMN-compliant CDP

Enterprise BPMN 2.0-compliant process engines generally support standard computer scripting languages such as Javascript, Groovy, and Python, for use with the script task. For specific applications such as the compliance audit in a particular domain, however, it may be necessary to develop a domain-specific BPMN-compliant process engine that incorporates a purpose-built domain specific language. This may be necessary for handling certain concepts and types of information that are specific to the domain. Exemplar domain-specific languages that have been used in conjunction with BPMN-compliant CDP include the high-level query language RKQL (Regulatory Knowledge Query Language) [13] and BIMRL (BIM Rule Language) [15].

4 Legal Knowledge Model (LKM)

There are two aspects to the legal knowledge conveyed by legal documents, namely the document structure and text (or the literal content) and the semantics (or the logical content). It is considered essential for any computerized legal document to provide separate representations of the literal and logical content of the document that allows them to be maintained coherently. The literal content representation maintains user familiarity with the look and feel of the original text, which would promote its adoption in

practice. The logical content representation is primarily intended for machines to process, but also allows human experts to manage and maintain. A Legal Knowledge Model (LKM) pertains to both representations operating together as one entity.

4.1 LegalDocML

LegalDocML [10] is a standardization of Akoma Ntoso (Architecture for Knowledge-Oriented Management of African Normative Texts using Open Standards and Ontology) [16], which can be used to represent the literal content of a legal document. The current version of the schema is Akoma Ntoso 3.0, which has gained popularity in the legal domain for representing legislative and judicial documents. LegalDocML is compatible with an existing European standard CEN MetaLex [17], which has been used to represent the entire set of Dutch regulations.

Among other features that have been developed using open standards, LegalDocML supports document workflow tracking with the ability to capture the entire lifecycle of a document, and provides an automatic version control and tight coupling with its LegalRuleML counterpart.

As it is exchanged in XML, any LegalDocML representation of a legal document can be rendered in user-friendly readable formats such as HTML or plain text while maintaining the structure and literal content of the source document. This allows the user to navigate a digital version of a legal document in the same way as they would do with the original paper-based document.

4.2 LegalRuleML

LegalRuleML [11] is inherited from RuleML [18] with extended features specific to the formalization of norms, guidelines, and legal reasoning. This emerging open standard can be used to represent the logical content of any legal document.

As the name implies, LegalRuleML enables normative provisions to be represented as rules. The encoding of norms into rules is currently a manual process. It is expected that the development of LegalRuleML representations of legal documents is to be undertaken by the same government agencies responsible for authoring the legal documents in the first place. However, NLP and other AI techniques may soon be able to assist in automating some of the encoding process.

Each rule represented in LegalRuleML has a unique key that is associated with its source provision in LegalDocML Any changes in the source text in LegalDocML would trigger the need to update its rule representation in LegalRuleML.

5 Worked Example

As a worked example, a common workflow that is used repeatedly throughout various stages of a building's life-cycle has been selected. This workflow checks if the opening direction of a door from any space is compliant with provisions in the fire regulations. For this example, the provisions of the New Zealand Building Code has been selected.

The is taken from Paragraph 3.2.6 "Direction of Opening" of the C/VM2 document [19], which stipulates that "*Doors…shall be hung to open in the direction of escape…These requirements need not apply where the number of occupants…using the door is no greater than 50.*"

In another word, this regulatory provision prescribes that any door used for exit by potentially more than 50 people from a space must swing open in the outward direction.

5.1 Space Activity and Occupant Load

The potential occupant load of a space can often be determined by the type of primary activity designated for the space. This is particularly relevant for spaces with the potential capacity of accommodating a large number of people. For example, the expected occupant load in a space where the crowd is normally standing as part of the activity would be higher than if the same space is furnished for a different activity with some seating arrangements such as couches or sofas.

The activity of a space may change multiple times throughout the lifecycle of the building. Consequently, the expected occupant load of the space also changes accordingly. Therefore, it is necessary for building designers and operators to audit building components in the space (such as the door and its opening direction) for compliance as changes occur in the designated use or activity type of the space.

In a conventional procedure, knowing the activity type of a space, one can look up the prescribed occupant load density (defined as the number of persons per unit floor area) from the building code or building regulations. Given the floor area of the space and the occupant load density, the potential occupant load can then be calculated.

5.2 Exemplar CDP

An exemplar CDP that represents a compliance path implied by the selected regulatory provision given in this worked example is shown in Fig. 3. This workflow describes the procedure of checking the opening direction of a door from a space based on the occupant load of the space for compliance with the specified requirement.
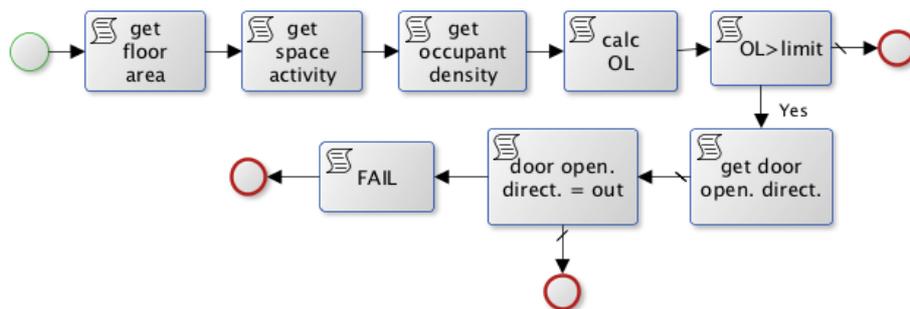

Fig. 3: CDP to check compliant door opening direction

In this CDP, it is assumed that the occupant load of the space is not available from the building model (PDM/BIM), but the type of activity designated for the space is known. Otherwise, the information may be provided via supplementary manual input.

The CDP starts by retrieving the floor area of a space from PDM/BIM. This is followed by a sequence of tasks to retrieve the space activity type, look up the prescribed occupant load density based on the activity, and then calculate the occupant load. The next task is to look up the required upper limit that shall not be exceeded from LKM (i.e. 50 in this case) and check if the calculated occupant load is within this threshold. If the occupant load is less than 50, then no further check is required and the audit process is deemed completed. Otherwise, the next task is to the check the opening direction of the door of the space and decide if it passes or fails the audit.

An exemplar instruction (in RKQL script) embedded into the task of calculating the occupant load (OL) is shown below as the BPMN-compliant XML representation of the script task (i.e. an excerpt of the CDP given in Fig.3):

```
<scriptTask id="PO_p1037">
  <incoming>PO_p1036</incoming>
  <outgoing>PO_p1043</outgoing>
  <script>SET OL = (floorArea/occupantDensity)</script>
</scriptTask>
```

5.3 Information from PDM/BIM

Provided that the building model contains a space (*IfcSpace* object), it is very likely that it also contains the floor area information, which is one of its defined quantity. The other piece of information expected from the PDM/BIM for this worked example is the space activity designation, which may or may not be present in the model, but it can be made available through supplementary input either directly into the PDM/BIM or supplied separately through the CDP as part of the process.

In the IFC4 specification (also in IFC2x3 version), the space activity designation is conveyed by *OccupancyType* of the *Pset_SpaceOccupancyRequirements* property set related to the *IfcSpace* object. This property set also includes *OccupancyNumber* and *AreaPerOccupant*, which can both be used to convey the information needed for the compliance audit process, if required. The floor area of a space is specified in IFC4 schema by *NetFloorArea* of the *Qto_SpaceBaseQuantities* quantity set related to the *IfcSpace* object.

5.4 Information from LKM

In accordance with the instruction in the CDP, the process engine would extract the occupant load density associated with the activity type from the LKM, if it was not specified in the PDM/BIM. This exemplar regulatory provision is obligatory in nature and can be represented in LegalRuleML as shown in the excerpt below:

```
<lrml:PrescriptiveStatement key="3.2.6_R1.K01">
  <ruleml:Rule>
    <ruleml:if>
      <ruleml:Atom>
        <ruleml:Rel>greater than</ruleml:Rel>
        <ruleml:Var>occupantLoad</ruleml:Var>
        <ruleml:Data>50</ruleml:Data>
      </ruleml:Atom>
    </ruleml:if>
    <ruleml:then>
      <lrml:Obligation>
        <ruleml:Atom>
          <ruleml:Rel>doorOpenDirection</ruleml:Rel>
          <ruleml:Con>out</ruleml:Con>
        </ruleml:Atom>
      </lrml:Obligation>
    </ruleml:then>
  </ruleml:Rule>
</lrml:PrescriptiveStatement>
```

The rule representation contains one condition, i.e. the occupant load of the space under audit is greater than 50. Otherwise, the rule is not applicable. If the condition is true, then the consequent of the rule is that the opening mode of the door must be in the outwards direction to pass the audit process. Otherwise, it would fail the audit process.

In the LegalRuleML excerpt above, the rule has a unique key of "3.2.6_R1.K01", which is associated with the same key in the source LegalDocML representing the original text of Paragraph 3.2.6 in the C/VM2 document. This object relationship feature enables version control and tracking of any amendment in the source text, which would prompt an update of the rule representation, as necessary.

6   Conclusion

This paper has described a practical approach of automating compliance audit processes with a simple example to demonstrate how an open standard process model can be used to specify how to access and process the required information. Conventional audit processes are laborious and error-prone because humans are tasked with accessing and processing information repetitively, which is an activity that can be performed much more efficiently and accurately by machines.

Automation is possible when computer representations of the product and legal knowledge are available. However, human experts should retain their roles in designing the compliance audit process (CDP) and give instructions as neceesary, thereby providing a guidance to the automation process. As each CDP is a formal documentation of a design procedure, it can be audited and validated for correctness. This would minimise any human error in the process, especially when the CDP is intended to be used multiple times across different projects.

References

1. Nana, G.: Assessment of the Economic Impact of Efficiency Improvements in Building and Construction. Business and Economic Research Limited, Wellington, New Zealand (2003).
2. Dimyadi, J., Amor, R.: Automated Building Code Compliance Checking - Where is it at? In: 19th International CIB World Building Congress 2013. pp. 172–185. Queensland Institute of Technology, Brisbane, Australia (2013).
3. Eastman, C., Lee, J., Jeong, Y., Lee, J.: Automatic rule-based checking of building designs. J. Autom. Constr. 18, 1011–1033 (2009).
4. Ding, L., Drogemuller, R., Rosenman, M.A., Marchant, D., Gero, J.: Automating code checking for building designs - DesignCheck. Proceedings of the CRC for Construction Innovation International Conference. pp. 1–16. Gold Coast, Australia (2006).
5. Khemlani, L.: CORENET e-PlanCheck: Singapore's automated code checking system, http://aecbytes.com/buildingthefuture/2005/CORENETePlanCheck.html.
6. Dragoni, M., Villata, S., Rizzi, W., Governatori, G.: Combining NLP Approaches for Rule Extraction from Legal Documents. In: Proceedings of the workshop on "Mining and Reasoning with Legal Texts" collocated at the 29th International Conference on Legal Knowledge and Information Systems. pp. 1–13 (2016).
7. Zhou, Z.D., Ai, Q.S., Liu, Q., Yang, W.Z., Xie, S.Q.: A STEP-compliant product data model for injection moulding products. Int. J. Prod. Res. 47, 4497–4520 (2009).
8. ISO 16739: ISO 16739:2013 Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries. International Organization for Standardization, Geneva, Switzerland (2013).
9. Hietanen, J., Lehtinen, S.: MVD AND SIMPLEBIM (whitepaper), (2014).
10. OASIS: OASIS LegalDocumentML (LegalDocML) TC, https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legaldocml.
11. Palmirani, M., Governatori, G., Rotolo, A.: LegalRuleML: XML-based rules and norms. Rule-Based Model. Comput. Semant. Web. 298–312 (2011).
12. Dimyadi, J., Amor, R., Spearpoint, M.: Using BIM to Support Simulation of Compliant Building Evacuation. Proceedings of the 11th ECPPM. pp. 511–518. Limassol, Cyprus (2016).
13. Dimyadi, J., Clifton, C., Spearpoint, M., Amor, R.: Computerizing Regulatory Knowledge for Building Engineering Design. J. Comput. Civ. Eng. C4016001, 1–13 (2016).
14. ISO/IEC 19510: ISO/IEC 19510:2013 Information Technology - Object Management Group Business Process Model and Notation., Geneva, Switzerland (2013).
15. Dimyadi, J., Solihin, W., Eastman, C., Amor, R.: Integrating the BIM Rule Language into Compliant Design Audit Processes. Proceedings of the 33th CIB W78 International Conference 2016. pp. 1–10. Brisbane, Australia (2016).
16. Palmirani, M., Vitali, F.: Chapter 6 Akoma-Ntoso for Legal Documents. In: Legislative XML for the Semantic Web. pp. 75–100 (2011).
17. Boer, A., Winkels, R., Vitali, F.: Metalex XML and the Legal Knowledge Interchange Format. Springer-Verlag (2008).
18. Boley, H., Paschke, A., Shafiq, O.: RuleML 1.0: The overarching specification of web rules. In: Alferes, J.J., Bertossi, L., Governatori, G., Fodor, P., and Roman, D. (eds.) International Workshop on Rules and Rule Markup Languages for the Semantic Web (RuleML 2010). pp. 162–178. Springer, Washington, D.C., USA (2010).
19. MBIE: C/VM2 Verification Method: Framework for Fire Safety Design For New Zealand Building Code Clauses C1-C6 Protection from Fire. The Ministry of Business, Innovation and Employment, Wellington, New Zealand (2014).