Legal Compliance by Design (LCbD) and through Design (LCtD): Preliminary Survey

Pompeu Casanovas\textsuperscript{abc}, Jorge González-Conejero\textsuperscript{c}, Louis de Koker\textsuperscript{ab}

\textsuperscript{a}La Trobe Law School, La Trobe University, Melbourne, Australia
\textsuperscript{b}Data to Decisions Cooperative Research Centre, Australia
\textsuperscript{c}Institute of Law and Technology, Autonomous University of Barcelona, Spain

Abstract. The purpose of this paper is twofold: (i) carrying out a preliminary survey of the literature and research projects on Compliance by Design (CbD); and (ii) clarifying the double process of (a) extending business managing techniques to other regulatory fields, and (b) converging trends in legal theory, legal technology and Artificial Intelligence. The paper highlights the connections and differences we found across different domains and proposals. We distinguish three different policy-driven types of CbD: (i) business, (ii) regulatory, (iii) and legal. The recent deployment of ethical views, and the implementation of general principles of privacy and data protection lead to the conclusion that, in order to appropriately define legal compliance, Compliance through Design (CtD) should be differentiated from CbD.

Keywords. Compliance Survey, Compliance methodologies, Compliance by Design, Legal Compliance through Design, Regulatory Compliance

1. Introduction

Compliance —and particularly legal compliance— is a popular topic, featuring for example in research projects on big data analysis, blockchain and other digital ledger technologies, digital currencies, fintech, regtech, crowdsourcing, tax regulations, smart cities, cloud computing, normative Multi-Agent Systems, electronic institutions, health, security, data protection, and privacy.\textsuperscript{1} It is the subject research matter of several EU H2020 Projects.\textsuperscript{2} Previous projects — especially COMPAS\textsuperscript{3}, OPENLAWS\textsuperscript{4}, EU Cases\textsuperscript{5} and BO-ECLI\textsuperscript{6} — developed conceptual toolkits. Several surveys on regulatory compliance have been already performed in the last ten years, including a meta-analysis

\textsuperscript{1} This paper is based on D2D CRC Deliverable \textit{Technical Bases for Compliance by Design (CbD)}. While the support of the Data to Decisions Cooperative Research Centre is acknowledged, the views expressed in this article do not necessarily reflect the views of the Centre.
\textsuperscript{2} Cf. especially TRUST (\url{www.trust-project.eu/}), e-COMPLIANCE (\url{http://www.e-compliance-project.eu/}), MIREL (\url{http://www.mirelproject.eu/}), and LYNX.
\textsuperscript{3} \url{http://cordis.europa.eu/fp7/ict/ssai/docs/finalreport-compas.pdf}
\textsuperscript{4} \url{https://info.openlaws.com/openlaws-eu/}
\textsuperscript{5} \url{http://eucases.eu/start.html}
\textsuperscript{6} \url{http://bo-ecli.eu/}
of peer-reviewed systematic literature reviews on business process compliance [1]. We have reviewed 280 publications so far. This paper introduces a scheme to perform a systematic survey in the immediate future [2], focusing on the dimension of legal enforcement and public law. The application of technology in relation to compliance is going far beyond the business and corporate fields in which the idea of Compliance by Design (CbD) was originally coined.

Different expressions and approaches can be found in the current literature on compliance, according to different fields and purposes, with different meanings — mainly Compliance (C), Regulatory Compliance (RC), Compliance by Detection (CbDt), Compliance by Design (CbD), and Legal Compliance by Design (LCbD). The Cambridge dictionary equates ‘compliance’ with ‘obedience’: “the act of obeying an order, rule, or request”. Broadly, compliance can be understood as conformity in fulfilling requirements, or demonstrating conformity with regulatory constraints. RC is denoting a previously selected set of requirements. This set can be also defined in many ways. E.g. ISOs point at conformance of business procedures and processes with laws, regulations, standards, best practices, or similar requirements.

An increasing number of aspects of compliance can be automated to minimise risks, save resources, and increase management security, efficiency and effectivity. Compliance with formal rules —i.e. rule compliance— requires the definition of a language, scope, and information processing specifications. Some compliance systems have already been patented [3] [4].

Compliance by Detection (CbDt) entails a conformity check during or after the runtime stage (in the execution environment). Therefore, if noncompliant behaviour with a set of rules is detected, the business process needs to be redesigned. Conversely, Compliance by Design (CbD) means that the set of rules is taken into account in the design stage of the business process. The conformity check takes place in advance, before and within the runtime stage. This approach has the advantages that: (i) a subsequent proof and corrections of compliance are not required; (ii) the approach is flexible as the generation can be repeated when rules are added, removed or changed; (iii) compliance is not only detected, but actually enforced [5]. Hence, CbD has a preventive side: it means that compliance “should be embedded into the business practice, rather than be seen as a distinct activity” [16].

Legal CbD (LCbD) is a term that was introduced to focus on the legality of the whole business process, mainly after the enactment of the Sarbanes-Oxley Act (2002), a US federal law that expanded and created new requirements for all public company boards and accounting firms. Actually one of the first uses of the term CbD in computer science occurred having LCbD in mind [29]. The term Holistic Compliance is also used to point out that LCbD “stands in contrast to simply complying with the rules, and thus,

7 https://dictionary.cambridge.org/dictionary/english/compliance
8 According to ISO/IEC 27002: “The organization must identify and document its obligations to external authorities and other third parties in relation to information security, including intellectual property, [business] records, privacy/personally identifiable information and cryptography”.
9 “When”, “how”, and “what” matter here. CbD and CbDt can be applied to the same system at different stages. The application of both approaches are usually recommended. The former during the design of the business process and the later during the running stage.
10 There were other interesting uses, e.g. [59], a civil engineering thesis on roundabouts stated: “Roundabouts encourage speed compliance by design instead of regulatory enforcement” (2004).
imperatively requires an integrated design of both relevant elements and the relationships amongst these" [6]. Regulatory Compliance is also a common notion, often referred in a broad way to denote “the act and process of ensuring adherence to laws” and involving analysing, checking, enforcing, and “discovering, extracting and representing different requirements from laws and regulations that affect a business process” [1].

Thus, we may distinguish between LCbD approaches focused on business processes — e.g. through goal-oriented modelling [7] — and those focused on legal knowledge, i.e. on requirements based on the properties of normative and legal systems (such as hierarchy, consistency, etc.) [8]. The definition of legal (not only documentary) sources to select and define requirements deserves some attention. We will suggest the notion of Compliance through Design (CtD) to explicitly encompass the social and institutional aspects that are not explicitly included by the regular way of approaching this subject (i.e. legal interpretation processes — beyond the conversations between experts and computer scientists —, institutionalisation, the interface between modelling and coordination, and the relation between citizens and the law).

In business studies, both CbDt and CbD are focused on regulations and legal texts, standards, outer policies and inner policies. In this scenario, we can take four different methodological stances to describe how they have been modelled: (i) deontic logic, (ii) Petri nets, (iii) graph-based tools, (iv) goal-oriented languages, (v) and semantics.

The remainder of this paper will be organised as follows. We will briefly summarise (i) business approaches to compliance, (ii) semantic languages, (iii) the recent synergy between legal theory, business languages, AI, and Normative Agent Systems (norMAS), (iv) and compliance methodologies. Finally, in the last section, we will briefly introduce the notion of CtD. Figure 1 shows the double circuit of sources and formalisation, following the main technical trends in business process modelling. It will be the basis to gear the comparative in this area.

![Figure 1](image-url)

*Figure 1. Survey analytical scheme (relation between sources and BC formalisation)*
2. Business process modelling

2.1. Graph-based Business Process Modelling

There are many business analyses in the literature about the way companies work and define business processes. It is a common trend to visualise business processes in a flow-chart format. However, this creates a technical gap between the format of the initial design of business processes, and the format of the languages that will execute these business processes. This gap needs to be bridged with a formal mechanism that maps the appropriate visualisation to the appropriate execution format. Therefore, there is a number of graph-based business process modelling languages, for instance Business Process Modeling Notation (BPMN 2.0 Specification, 2016), Event-driven Process Chain (EPC diagrams), Unified Modeling Language (UML diagrams), and UML activity diagrams (AD) [9] [10].

2.2. Business Process Model and Notation (BPMN)

The Business Process Model and Notation (BPMN) is a graphic representation of business processes in a business process model. The Business Process Management Initiative (BPMI) developed BPMN (which has been maintained by the Object Management Group (OMG) since the two organizations merged in 2005). BPMN is a standard promoted by OMG in order to ease the designing of business process models within different scenarios. The primary goal of BPMN is to provide a notation that is readily understandable by all business stakeholders - business analysts, technical developers responsible for implementing the technology, and business people who will manage and monitor these processes. Another goal is to ensure that XML languages designed for the execution of business processes can be visualised with a business-oriented notation [11].

2.3. Business Process Modelling and Notation – Query (BPMN-Q)

The Business Process Model Notation Query (BPMN-Q) is a visual language to query repositories of business process models. This language has been used in different scenarios where reuse or retrieval of process models based on a query pattern was needed. BPMN-Q query is represented as a business process diagram using the BPMN that are augmented with querying-specific constructs [12].

2.4. Temporal Deontic Logic and Computational Tree Logic

Deontic logic (DL) is a logic for representing and reasoning about concepts such as obligation, permission, prohibition and waived obligation. Various axiomatizations of DL have been proposed with different extensions: standard DL (SDL), Computational Tree Logic, and DL based on Event Calculus formalism. Several non-standard DL approaches have been proposed recently [13], and are experiencing a fast grow due to

Semantic Web and Normative Multi-agent Systems developments. Computational Tree Logic (CTL) is a language to express properties for model checking. In addition to logical operators and atomic propositions, CTL provides temporal operators to express properties that must hold through a number of states that are temporally related. PEN EL OPE [Process ENtailment from the Elicitation of Obligations and Permissions] [14] is a language to express temporal deontic assignments. It is mainly designed to generate compliant control-flow-based process models from a rule set of permissions and obligations.

2.5. Petri Nets

Petri Nets — or place/transition (P/T) nets — combine a simple graphical representation with a mathematical basis [15]. They can express locality of actions in distributed systems and allow modelling the lifecycle of several business processes, yielding a more compact model compared to explicit state machines. A CbD approach based on artefact-centric business processes is introduced in [5].

2.6. LegalRuleML

LegalRuleML is an effort to create a standard13 for the representation of norms, assuming the equivalence between temporal and defeasible logic, and eventually a general legal unified framework for rule interchange languages [16] [17]. LegalRuleML documents consist of metadata, statements (rules), and contexts. LegalRuleML is deemed to represent the particularities of the legal normative rules with an articulated and meaningful mark-up language, encompassing the following features: (i) defeasibility of rules and defeasible logic; (ii) deontic operators (e.g., obligations, permissions, prohibitions, rights); (iii) semantic management of negation; (iv) temporal management of rules and temporality in rules; (v) classification of norms (i.e., constitutive, prescriptive); (vi) jurisdiction of norms; (vii) isomorphism between rules and natural language normative provisions; and (viii) identification of parts of the norms (e.g., bearer, conditions); (ix) authorial tracking of rules [17]. Compliance in business processes and exchanges can be represented and are addressed stemming from this language of representation [18].

2.7. Ontologies

In the compliance framework, ontologies are used to represent the extracted knowledge from different legal and normative texts in a machine-readable format. Furthermore, XML Schema allows the execution of a reasoner algorithm against the ontology structure to determine the compliance status of sensitive assets. Focusing on compliance, and stemming from existing practices and methods, for example, the Compliance Management Ontology [CoMOn] [19] proposes a shared conceptualization. It is based on several core concepts (further linked and organised into two more tiers): business process management, culture management, obligations, programme, resources, risk management, and solutions.

13 [https://www.oasis-open.org/committees/legalruleml.charter.php](https://www.oasis-open.org/committees/legalruleml.charter.php)
3. Methodologies and Corporate Governance Models

Methodologies are related to the broad corporate governance models and standards that have been developed in the last twenty-five years. There are, for example, several ISO standards related to corporate and regulatory compliance and security, and several corporate governance models. Some models for IT Governance stem from COSO, COBIT, ISO 27002 (ISO 17799) and ISO 38500. There is some confusion around the different models, as they are meant to pursue different objectives that are not always compatible: (i) stewardship of IT resources on behalf of various stakeholders, (ii) planning, organizing, and monitoring the use of IT resources; (iii) creating value for the stakeholders; (iv) complying with national and international laws to avoid regulatory risks; (v) both protecting consumers and customising consumer experiences; and (vi) improving market quality.

3.1. ISO/IEC Standards


ISO 17799 (developed today by ISO 27001/02) represents a guide for implementing a set of policies, practices and procedures in order to consolidate the information security administered by an organization. ISO/IEC 27002 requires that management systematically examines the organization’s information security risks, taking account of the threats, vulnerabilities and impacts. Clause 6.1.3 of ISO/27001:2013 describes how an organisation can respond to risks with a risk treatment plan; an important part of this is choosing appropriate controls.

ISO/IEC 27002 seeks the preservation of (i) confidentiality (information accessible only to those authorized to have access), (ii) integrity (accuracy and completeness of information and processing methods), (iii) and availability (authorized users have access to information and associated assets when required). Section 18 points at compliance:

18.1 Compliance with legal and contractual requirements: “The organization must identify and document its obligations to external authorities and other third parties in relation to information security, including intellectual property, [business] records, privacy/personally identifiable information and cryptography”.

18.2 Information security reviews: “The organization’s information security arrangements should be independently reviewed (audited) and reported to management. Managers should also

---

14 This standard is based on the AS 8015-2005 Australian Standard for Corporate Governance of Information and Communication Technology (2005).
routinely review employees’ and systems’ compliance with security policies, procedures etc. and initiate corrective actions where necessary.”

ISO/IEC 29382, Corporate Governance of Information and Communication Technology, officially named ISO/IEC 38500 in April 2008, could be added to this list. It is intended to provide guiding principles to any organization, regardless of size or sector.17

3.2. Automated, Semi-automated, Inner, and Outer Compliance

Governance models and ISOS tend to overlap. Mapping COBIT, COSO and other audit IT models into ISO/IEC 27001/2 (ISO 17799) and ISO 38500 is now a common endeavour [20]. Researchers have found empirical evidence to assert that a firms’ effectiveness of IT steering committee-driven IT governance initiatives positively relate to the level of their IT-related capabilities. Thus, there is a relationship between IT-related capabilities and internal process-level performance. Improvement in internal process-level performance will be positively related to improvement in customer service and firm-level performance [21].

From the outset automation of security requirements has been one of the objectives of IT corporate governance. It belongs to the Security Requirements Engineering (SRE) domain, which falls outside the strict boundaries of corporate governance. The SQUARE methodology [Security Quality Requirements Engineering] developed at Carnegie Mellon University in 2005 [22], provides a step-by-step process for “eliciting and prioritizing” security requirements into software and systems early in the life cycle. The SQUARE methodology has been recently extended [23], making it useful as a framework. L-SQUARE addresses legal compliance when developing software, engineering systems or acquiring software. Complete SRE updated surveys do exist [24], and several ontologies have been already built, advancing a meta-model for knowledge reusing in security requirements engineering [25].

4. Legal Compliance by Design (LCbD)

In this Section, we will describe some trends that brings closer alignment to business and legal modelling. We will contend however that finding synergies is a more effective strategy than making transplants. Some methods and tools to restructure and adapt legal norms, policies and rights to Linked Open Data domains are also referred.

4.1. Regorous

Following the Regorous approach, compliance “is a relationship between two sets of specifications: alignment of formal specifications for business processes and formal specifications for prescriptive (legal) documents” [26]. The authors view the compliance ecosystem as the complete lifecycle between domain experts and modellers, including knowledge acquisition, checking, translation, and monitoring. Hence, modellers behave proactively and intervene and join in the description of needs, roles and tasks, annotating and enriching the required business information modelling, for as “compliance is not just

17 http://www.38500.org/
about the tasks to be executed in a process but also on what the tasks do, the way they change the data and the state of the artefacts related to the process, and the resources linked to it” [ibid.]. Accordingly, process models must be enriched with such information.

To address the problem of how to get data for data control tags and where to get it from, a query-based methodology to annotate process models is advanced. This is called the “Regorous architecture”, extended with the semantics of the LegalRuleML model.

The authors sequence the Regorous steps as follows: 
1. generate an execution trace of the process; 
2. traverse the trace: for each task in the trace, cumulate the effects of the task using an update semantics; 
3. use the set of cumulated effects to determine which obligations enter into force at the current tasks, by calling a reasoner; 
4. add the obligations obtained from the previous step to the set of obligations carried over from the previous task; 
5. determine which obligations have been fulfilled, violated, or are pending; 
6. if they are violated, check whether they have been compensated; 
7. repeat for all traces. “In conclusion, a process is evaluated as compliant if and only if all traces are compliant (all obligations have been fulfilled or if violated they have been compensated), or it is evaluated as weakly compliant if there is at least one trace that is compliant” [18].

It should be noted that there is a great bulk of descriptive work carried out before, during, and after modelling. In essence, legal compliance in terms of this approach is compliance with the extracted conceptual legal model.

4.2. Mercury

Cooperation between Subject Matter Experts (SME) and modelers as Semantic Technology Experts (STE) is also specifically addressed in the Mercury approach [27][28]. Mercury is a linguistic tool pointing at legal extraction and interpretation to reach compliance (e.g. for financial and tax purposes). It is composed of a structured English vocabulary based on SBVR and represented in a XML Schema capturing rulebook and vocabulary entries. The rulebook contains regulative and constitutive rules; the vocabulary encompasses the actions and factors that determine a rule’s applicability and its legal effect. Thus, Mercury focuses on the extension of linguistic terms complemented with a regulatory interpretive methodology (RIM): “Mercury represents rule statements contained in regulations and describes the concepts used in those rules in a terminological dictionary” [28]. It leans on SBVR, but introduces a different way to manage the logical formulation of SBVR rules, aligning them with a consistent legal interpretation. This is a “constructive” trend, representing legal knowledge beforehand as a bridging tool between SBVR and its effective implementation.

4.3. Eunomos

Eunomos is a legal knowledge and document management system focused on identifying (i) norms, (ii) related norms, (iii) legislative modifications, (iv) different interpretations of the same norms. It “focuses on identifying norms, cross-references and semantic similarities, with a clear structure for representing multiple interpretations and normative change.” [30][31]. The architecture of the system is composed of three levels:

---

18 The key aspects of the methodology are: (1) to enrich business process models with semantic annotations, (2) to extract control objectives from business rules, and (3) to formalise the control objectives in an appropriate logical formalism.
(i) a legal document management system (a database of norms in legislative XML, a database collecting the network of references between laws, and a database classifying single articles or items of legislations in different domains); (ii) a legal knowledge management system (a database of concepts and of relations connecting them, together with the terms associated to concepts, plus a legal document management system to associate concepts and articles or items of legislations, and a database of prescriptions; (iii) an external tier (a database of user profiles for login purposes and for keeping information about their domains of interest, dispatching alerts concerning updates etc.). Databases are populated by web spiders collecting daily new legislation, identified by their URN identifier obtained by translating the human language title of the law.19 The database architecture is divided into two independent parts: the Legal Taxonomy Sylabus ontology, and the legal texts repository.

It is worth noting that compliance, diversity of legal interpretations, and attention to stakeholders (lawyers, officers, and citizens) are at the core of the Eunomos focus. Hence, Eunomos is also conceived as a commercial web service—called Menslegis—for business compliance professionals. It encompasses tracking legislative changes (through Cosine Text Similarity), enduring traceability from the text of laws and their application to business processes, and a workflow for designing compliant business processes [32].

4.4. Legal-URN and Legal Goal-Oriented Requirement Language (Legal GRL)

Eunomos is a legal knowledge management system, using lightweight ontologies, annotations, and created concepts to link legal documents with end-users’ needs. Legal-URN (URN stands for User Requirements Notation) and Legal GRL focus on a preliminary stage and on the dimension of Requirement Engineering techniques to carry out business process modelling and business compliance. A meta-model to assemble and make compatible both systems has already been proposed [33]. The first one aiming at legal knowledge management; the latter focusing on goal oriented modelling.

URN constitutes an IT standard since 2008, as it is a language that visually supports the elicitation, analysis, specification, and validation of requirements. Legal URN is the extension of a requirements management framework for business process compliance, connecting organisations and regulations within three different layers: (i) documentary, (ii) teleological (organizational and legal goals), (iii) business processes models.20 To develop Legal GRL models from regulations, the Hohfeldian jural classification is fleshed out and refined for computational objectives. Duty-Claim, Privilege-NoClaim, Power-Liability and Immunity-Disability are turned up into a format that exploits deontic modalities at the GRL level. This occurs in two steps: (i) classifying each statement of the legal document based on Hohfeld’s classes of rights and annotate it accordingly

19 “Then the text of the norm is automatically translated into legislative XML using a parser. References in the text of norms, already tagged in XML, are collected in a database. Then norms are classified semi-automatically, and the collection of concepts can start. The system is implemented in PHP for the web application, Javascript and Ajax, for the front end, XML and XSLT for the documents, and C++ for the web spiders retrieving legislations” [31].

20 “GRL is tailored here through a lightweight legal profile to capture the elements of law, including deontic modalities. This new profile is called Legal GRL. Legal GRL is part of the Legal-URN framework, which includes Legal UCM as well [Use Case Maps (enabling to capture business processes)]” [34].
(identifying “subject”, “verb”, “actions”, “preconditions”, “exceptions” and “cross-references” parts in each legal statement); (ii) refining the Hohfeldian classes of rights into deontic Permission and Obligation goals, developing the goal model of the law, and annotating the intentional elements with “Permission”, “Obligation”, “Precondition”, “Exception”, and “XRef” stereotypes [34].

4.5. Nòmos

Legal GRL and Nòmos share the use of i* modelling language (i-star, distributed intentionality), and their interest for the Hohfeld jural squares as a semantic solution to model legal relations. In Nòmos, intentional compliance is defined as “the assignment of responsibilities to actors such that, if every actor fulfils its goals, then actual compliance is achieved” [35] [36]. Thus, it intends finding a bridge between two different sets of concepts (law / requirements) using normative propositions as atomic elements and jural relations, actors, and actions as components: Laws are about rights and obligations, privileges and liabilities; requirements, are about stakeholders’ goals and system behaviours to meet them. This analysis (Nòmos 2) has lately been extended to the distribution of responsibilities stemming from social and legal roles (Nòmos 3) [37].

4.6. Rights Expression Languages (REL)

Rights Expression Languages (REL) are computer languages also created to handle and manage rights and obligations (permissions and prohibition) about content use, i.e. a “format for describing rights, i.e. permissions and constraints, related to the use of content.” [21]. REL may use Entity-Attribute-Value model, as for RDF, to structure its description of rights as a list of (i) entities (such as a “work”, or “asset” for a license), (ii) attributes (properties, such as actions that are permitted or forbidden, as constraints), (iii) values for these properties, from a pre-defined vocabulary, i.e. using or modifying the work). Well-known REL are ccREL (Creative Commons language to express their licenses) [22]. XrML, the eXtensible Rights Markup Language which has also been standardised as the Rights Expression Language (REL) for MPEG-21 [23], and W3C Open Digital Rights Language (ODRL).

Quite recently, in February 2017, the W3C Permissions and Obligations Expression (POE) Working Group updated Working Drafts of the ODRL Information Model and ODRL Vocabulary & Expression. The ODRL Information Model “offers a framework for the underlying concepts, entities, and relationships that form the foundational basis for the semantics of ODRL expressions” [38]. The ODRL Vocabulary & Expression “describes the potential terms used in ODRL Policy expressions and how to serialise them”. The terms form part of the ODRL Ontology and formalise the semantics.

[22] https://www.w3.org/Submission/ccREL/
4.7. Ontology Design Patterns (ODP)

An ODP is a structural pattern, a reusable successful solution to a recurrent modelling problem. It brings together expert and engineering knowledge, and can be used to build other domain or core ontologies for a particular field of knowledge [39]. Focusing on Rights Expression Languages (REL), an Ontology Design Pattern (ODP) has already been delivered for linked data licenses (from six rights expressions and policy languages (ODRL, MPEG-21 REL, XACML, ccREL, MPEG-21 MVCO and WAC). "In particular, the core idea of the pattern is to model: a rights expression which allows/prohibits/obliges to make an Action (Right) to an Agent over a LD resource under a condition" [40]. ODP are semantic regulatory tools with deontic components. They can be built without an integrated legal architecture. Instead, they are based on extended vocabularies and a better definition of concepts behind them, structured through different kind of ontologies. Term banks (lexical datasets) may contain millions of words in many natural languages.

There are different methods for approaching this multilingual complexity. In Europe, e.g., with 28 official languages: (i) controlled vocabularies, implemented in terminology database (such as IATE run by all the main EU Institutions), (ii) thesauri (as EUROVOC), (iii) semantic lexicons or lightweight ontologies (as WordNet, EuroWordNet and, in the legal domain, JurWordNet, EuroVoc Thesaurus). Structural patterns such as Ontolex-Lemon work as standard to represent lexicons relative to ontologies, and they can be used to encode term banks as RDF [41].

However, a standardisation of ontology reuse practices is still missing [42]. Linked data resources are being introduced to facilitate it, at least partially. E.g. Framester is a large RDF knowledge graph (currently including about 30 million RDF triples) acting as a hub between FrameNet, WordNet, VerbNet, BabelNet, Predicate Matrix, etc. Framester aims at leveraging “this wealth of links to create an interoperable and homogeneous predicate space represented in a formal rendering of frame semantics” [43].

5. Compliance through Design (CtD)

5.1. Limits of Compliance Languages

Compliance is at the crossroads between the linguistic way of conceiving legal components as requirements, and business modelling. As stated, the business process and task modelling constitute the first step, a well-trodden path in the specialised literature right now. The key question is how compliance requirements can be formally specified to enable the application of automatic analysis and reasoning technique for their verification [44].

To find an answer, the COMPAS authors performed a comparative analysis of compliance request languages, based on the examination of a wide range of compliance requirements and relevant frameworks such as Basel II, Sarbanes-Oxley, IFRS, FINRA (NASD/SEC), COSO, COBIT and OCEG. They compared (i) Linear Temporal Logic (LTL), (ii) Computational Tree Logic (CTL), (iii) Formal Contract Language (FCL), (iv) Metrical Temporal Logic (MTL), (v) Timed Computable Tree Logic (TCTL).

24 http://ontologydesignpatterns.org/wiki/Submissions:LicenseLinkedDataResources
25 See the core-Lemon structure at http://lemon-model.net/lemon-cookbook/node3.html
Although they found a high level of expressiveness, not all relevant rules in a real setting scenario (use case) could be represented. Their conclusion is that “decision on the use of a particular formal language is context-dependent and should be based on the nature, complexity and source of compliance requirements” [44]. They have recently introduced a compliance-request language meta-model for applying compliance patterns through Compliance Request Language (CRL) [45]. CRL is formally grounded on temporal logic and enables the abstract pattern-based specification of most compliance requirements.

5.2. Compliance in Normative Multi-Agent Systems [norMAS]

Sociotechnical systems involve a combination of software systems, people, and organizations. Norms can be broadly understood as a social device to allocate obligations and rights. Compliance with social norms related to agency and coordination of agents, is a classical problem. Legal compliance situates it in an institutional context, in which legal requirements have to be modelled to bring about acceptable results in social monitoring and auditing. The question here is not whether legal requirements can be automated, but how to define what is law (or what “count as” law). This is the problem of modelling audits for public services [46]. This case study entails implementing automated controls in the procurement process for public transport services for the elderly and disabled in the region of Eindhoven (care-related taxi services) [46].

The authors show that fully automated control might help to lower control costs, to prevent contractual misstatements, and to increase the quality of the audit. Nevertheless, interestingly, they don’t take a de jure approach: their research questions are focused on evidence, and related to issues such as transaction costs, auditing roles, and responsibilities. Thus, they use the term ‘compliance by design’ “in a broader sense, where the design of compliance measures encompasses an integrated view of organizational, procedural, and technical measures”. The authors acknowledge that tasks like data collection, monitoring, and triggering warnings can be automated, “but even in a fully automated control system, an auditor must assess the appropriateness of the design and verify operating effectiveness of the system as specified.”

This perspective is quite close to what we mean by CtD (Compliance through Design), i.e. the practical construction of integrated ecosystems, involving CbD, community-building, and the implementation of regulatory institutional designs alike, to increase the levels of transparency and public accountability. This concept will be unpacked in the next Section.

5.3. Legal Compliance through Design (LCtD)

Compliance surveys stressed that main areas for contributions have been in extracting legal requirements, modelling them with goal modelling languages, and integrating them with business processes [7]. The authors make clear that “compliance analysis is based on certification (that business processes comply with regulations) or auditing (ensuring the continuation of compliance)” [ibid.]. Other studies highlight (i) that “in practice, existing compliance-checking approaches have rarely been applied” [47], (ii) and “what is missing permanently except in the peak year 2009 is work geared towards compliance in the after execution phase” [48]. The recent meta-analysis of the literature (2016) confirms that—compared to relevant requirements, methodologies, and guidelines—very few attention has been payed to compliance enactment: “Compliance
analysis tasks and especially compliance enactment tasks have been neglected, as they are mentioned in only 16% and 4% of the referenced studies reviewed, respectively” [1].

Besides, beyond business processes, modelling law is not an easy task: “Just as courts must struggle to interpret the law when ambiguities are present, so must users, be they requirements engineers or developers, make crucial interpretation decisions during requirements gathering and software design” [8].

Especially at the crossroads of liberty and security — e.g. in constitutional principles, privacy and data protection — legal compliance enactment requires further information and conditions that can be modelled, but not hard-modelled [49] [50]. Computational requirements and social conditions as described in social sciences overlap but are not identical. Decisions have to be made at each stage of the modelling process. Conversely, legal conceptual models require selecting and using formal languages that are not meant to capture all relevant elements at the same time [51].

Still, artificial intelligence techniques enhance legal reasoning. Technology certainly eases the tasks of annotating, enriching, classifying, clustering and retrieving content, but also adds complexity to basic legal inferential operations such as interpretation, application, implementation, and enforcement. This means that there is a need both for representation languages and for intermediate regulatory models to anchor them into real settings and organisations. The assumption that laws are embedded in self-contained documents where their semantics can be automatically or semi-automatically extracted, applied, and eventually enforced does not hold if it is not complemented with regulatory instruments especially tailored to blend compliance systems with their human interfaces, uses and environments. It is our contention that this hybridation process deserves a closer attention.

This is not new, and there is a common awareness about the difficulties raised by algorithmic governance and linked open data analysis in this field. Moving from a natural language legal text to the respective set of machine-readable conditions is a real challenge, still under development [52]. “Public information chains” [53] and the notion of “near-compliance” [54] (applying strategies and tactics from the beginning, at the pre-conceptual modelling process) have already been advanced to cope with human-machine interfaces when privacy and public values, policies, and principles are at stake. The European Legislation Identifier (ELI) fosters citizens’ participation in law-making and regulations [60].

After having examined the state of the art in law and the web of data [55], and the regulation of big data [56], we would like to suggest the concept of Legal Compliance through Design (LCtD) to complement LCbD by recognizing the role of social, political, and economic conditions (as pre-conditions) and governance and ethical requirements (as constraints) when designing legal compliance, encompassing norms and principles that require a balancing of competing rights, obligations or policies.

LCbD should be complemented by LCtD in these cases, as legal implementation and compliance enactment is generally not only a technical issue but also a practical and theoretical one. The relation between different meta-models lies at the core of this approach: legal compliance cannot be comprehensively automated only by means of normative and linguistic tools. Implementing rights raises the problem of defining authority and democratic policies at the social and political level as well. Thus, ensuring basic compliance with the legal requirement is only one part of the complexity of designing an institutional compliance response. This brings about both advantages and

some risks, because the evolving social conditions of interpreting, fixing, and implementing the law should be taken into account and explicitly and separately addressed in specific contexts and environments. In short, legal sources (sources of authority), should not be confused with documentary resources. Sources require a previous theoretical identification to build and enact the regulatory model. These complexities can perhaps be better explained using an example:

Generally national laws (parliamentary statutes or regulation adopted under such statutes) require banks to identify and verify the identity of each customer by collecting a person’s name and other identifying particulars and, depending on risks, verifying some of all of these against government-issued documents or reliable data. That is what the law requires. The management of the bank on the other hand must decide how the bank should respond to that requirement. In cases where the national laws regarding customer identification and verification were prescriptive banks sometimes decided that they wish to go beyond the basic legal requirements and actually collect and verify more information about each customer [57]. In some cases they did so as they thought that what parliament requires is insufficient to mitigate their identity fraud risk. In other cases they asked more information because they wanted to profile the customer for marketing purposes [57]. Bankers looking at such a legal requirement may think: “This Act is 20 years old and predates mobile phones. If we interpret the requirements using the ‘spirit of the law’ interpretational approach, we believe that Parliament in the context of today would have required us to get the mobile number and that is why we require that from the customer”. Another bank may say: “It is costly to collect and store information and we will not ask mobile phone numbers as it is not required by the text of the Act” (i.e. a literal interpretation). Practical challenges may also dictate how the management of a bank would respond to such a legal requirement, for example where they may wish their institutions to collect more data on each customer, but their dated IT systems do not allow them to do that. They then have to settle for the time being for what is doable for them, planning to implement what is desirable when the system is overhauled. [57] Fashioning a CbD/CtD approach is therefore not only legal or regulatory compliance response to the general interpretation of a legal text, but rather a process to embed the compliance response elected by the particular institution.

It is furthermore important to appreciate that automated or semi-automated legal compliance is not simply compliance with the text of the requirement. It is instead compliance with the conceptual models elicited or extracted out of several sources through a knowledge-acquisition process that is not neutral, but a policy-driven set of tasks. For example, customer identification and verification practices can support or hinder a government policy of enhancing broader financial inclusion, especially of persons from socially marginalised communities. In our example, the management of a bank may or may not consider a public policy of financial inclusion when considering how the bank should respond to the legal requirement of customer identification and verification [58]. Their choice will affect the extracted regulatory model to be embedded in the design process.

Turning law into legal knowledge entails a previous step that has often been incorporated into running systems as a necessary assumption. But decisions and their underlying rationale should be elucidated and made explicit. The relevance of business decisions around compliance responses should not be ignored. In short, in more complex cases, especially where contending rights, obligations or policies are at play, a compliance response is not merely a response to a legal/regulatory requirement, but to a more complex set of variables that should be also elucidated and taken into account. This
is related to institutional building and strengthening. Recognising the concept of LCtD, and distinguishing it from LCbD, would enable designers to respond appropriately to these complexities, when present.


6. References


[28] Ceci, M., Butler, T., O’Brien, L. and Al Khalil, F., Legal Patterns for Different Constitutive Rules,. Mfrel Project, 2017


Proceedings of the 1st Workshop on Technologies for Regulatory Compliance


[59] Eickman, T.J. The evaluation of modern roundabouts as an alternative to signalized and two-way stop controlled intersections in a urban and rural environment (Doctoral dissertation, Montana State University-Bozeman, College of Engineering, 2004).