CREATe Working Paper 2014/12 (October 2014)

Self-enforcing or self-executing? What Computational Copyright can learn from LKIF Transaction Configurations for Eurobonds

Authors

Orlando Conetta
Pinsent Masons LLP
Orlando.Conetta@pinsentmasons.com

Burkhard Schafer
University of Edinburgh
B.Schafer@ed.ac.uk

A much shorter and more concise version of this report was published as Conetta, O., & Schafer, B. (2013). LKIF in Commercial Legal Practice: Transaction Configuration from Eurobonds to Copyright. In JURIX (pp. 63-72).

CREATe Working Paper Series DOI: 10.5281/zenodo.12432.

This release was supported by the RCUK funded *Centre for Copyright and New Business Models in the Creative Economy (CREATe)*, AHRC Grant Number AH/K000179/1.

KEYWORDS: Transaction Configuration, Transaction Framework, LKIF, CommonKADS, normative assessment.

ABSTRACT Uptake of Artificial Intelligence approaches in commercial practice has been low. By contrast, "Copyright by design" (DRM) and "Privacy by Design" have emerged as (commercially) successful applications of computer technology to legal issues, and are now closer to realizing the formalist ideal of a self-enforcing law than traditional approaches to legal AI and their attempt to model legal reasoning explicitly. DRM, and to a lesser extend PETs, have however also their detractors, causing commercial, societal and legal problems. This paper tries to rejoin the two approaches to computer technology in law, learning what can be learned from the success of DRM but trying to address its shortcomings by remaining firmly within the tradition of fully explicit legal modeling in the AI and Law tradition. For this, the paper presents a new theory, called Transaction Configuration, that tries to increase the practical utility of ontology driven approaches. It describes the main task common to contract lawyers in the performance of their work, and was developed in the course of a case study at a magic circle law firm in the City of London. Using Eurobond Transactions as a proof of concept, we discuss first how Transaction Configuration provides a practical context for legal normative assessment, second, we analyze the potential to extend this approach to complex forms of transactions, in particular copyright licensing. With DRM, it shares the idea that the best target for a computational approach to law is not the legal reasoning for a judge, but the automated enforcement/application of a contract or license. With traditional AI, it shares the emphasis on fully explicit modeling of legal reasoning.

1. INTRODUCTION

Adoption of Artificial Intelligence (AI) technologies in commercial legal practice remains very low. In a 2003 study of sixty six law firms, located in Australia and Norway, it was found that expert systems development scored close to the lowest level on their scale [1]. Beyond technological limitations Oskamp and Laurits postulate broader reasons for the absence, and attribute a major factor to a poor understanding of what lawyers want. They call for a two way dialogue between legal practice and AI research, writing:

> "until practicing lawyers see clear and immediate benefits of specific applications they will be reluctant to use them." [11 sec 4]

The position of copyright law in this assessment follows with to important exceptions a similar pattern. Until very recently, it was one of the more under-served fields of legal AI research. We analysed 15 years of contributions to the proceedings of ICAIL and JURIX as the main conferences that serve the AI and Law community, and all 21 years of the Journal for Artificial Intelligence and Law as the main journal in that field. We also carried out a keyword search via Google Scholar, DBLD and CiteSeer to

complement the survey.[1] Copyright emerged only recently as a topic of interest, and in particular did not play a role in the "classical" period of legal expert systems that model judicial reasoning.[2] In line with the general experience of limited take-up by legal practice, there was no evidence that AI technology had played a transformative role in copyright licensing or litigation, though both will have benefited like all fields of legal practice from "generic" computational tools for tasks such as information retrieval, case management or compliance assurance.

The first prominent contribution that we were able to find was the formal representation of an upper level ontology for copyright, which however constituted only one part of a larger investigation into formal models of Intellectual Property by Contissa and Laukyte in 2008 [22]. This paper reports mainly problems that were encountered in the process of developing a formal ontology for IP law. Generally, the emergence of formal ontologies as a building block of the semantic web has given a major boost to the research into formal representations of copyright law, and our own study follows this trajectory.

Formal ontologies play a central role for the semantic web. The prevalence of license statements attached to digital objects in the digital economy yielded the first exception to the general rule that research in AI and Law rarely leaves the confines of academic conferences, proof –of-concept prototypes and feasibility studies. However, users of the technology are not the traditional targets of legal AI research – law firms, courts and public administrations – but Internet publishers of all hues. Creative Commons for instance has made licenses available in RDF ("Resource Description Framework") format, which allows web publishers to embed license information in machine readable format in web pages, documents and mp3 files. RDF, first published in 1999 and substantially revised in 2014 is a specification for metadata data models issued by the World Wide Web Consortium (W3C). It can be used to describe conceptually information about resources that are implemented on the web, and in this way not only assist information retrieval, but also knowledge management and reasoning application. At its core philosophy is the idea of making statements about web resources in the form of subject-predicate-object expressions. This can be used to identify the author of a resource, or indeed its legal status ("This eBook has the property of being licensed through a CC-BY license")

In these triples, the subject (here: ebook) denotes the resource, the predicate denotes the type of relationship between the subject and the object (here: "being licensed though") and the object specifies the relationship (here" CC-BY license). This approach of labeling resources is the next step in the progressive evolution of the semantic web. It enables automated storage, exchange, and use of machine-readable information distributed throughout the Web. However, the simplicity and high degree of abstractness of the RDF data model also allows it to be used in applications outside semantic web activity, such as knowledge management.

Building on the notion of semantic web representations for digital resources and incorporating them into other knowledge management tasks this emerged as one of the most promising avenues for computational copyright law. An overview of these

[1] Search string (Boolean) ""copyright law" AND "artificial intelligence" OR "knowledge engineering" OR "knowledge representation""

[2] by contrast, patent law makes a much earlier, and much more prominent appearance, undoubtedly driven by the prevalence and importance of patent databases

approaches, and also an indication of the uptake this research, if not by lawyers then by web publishers, can be found in [24]

An important milestone in developing this approach further and reducing the gap between semantic web and knowledge management applicaitons was Gordon's paper at ICAIL 2011. Even though published only as a short paper, it has garnered in the relatively short period since its publication a considerable interest, with several follow up studies [20, 21]. Even more recently, Governatori et al [19] presented an approach to assist commercial exploitation of the Web of Data, with the aim of handling the licensing terms associated with data in an automated way. Conflicts between Creative common licenses in particular are a well known problem due the "viral" nature of the Share Alike provision, handling the resulting complexities a natural target for legal AI analysis.

Finally, a study by Palmirani et all [23] is worth mentioning. It discusses one example from copyright law to illustrate the role of extensions to XML standards for the semantic web. We discuss this application in some more detail, partly because it is typical for studies into web supported computational representations of legal norms, partly because the specific approach to legal AI and the formal languages that it choses, semantic representation of legal rules through a formal ontology, is very similar to our own. However, the methodology that is used to translate law into formal models is markedly different, as is the overall "philosophy" of developing computational tools for legal practice. Palmirani et al follow a top-down, highly generic and context independent approach that requires very little legal domain expertise, and as a result does not focus on any specific application or business case for legal AI. Our approach by contrast follows very closely legal practitioners in their everyday practice. It develops solutions that are much more closely tailored towards specific applications and domains, and require a much higher degree of expert input as a frontloaded development cost. Our two approaches are best understood "complementary", in the sense that Palmirani provide a generic computational framework, within which our own approach can the "grab" representations of legal concepts and automatically adopt them to the specificities of the situation which the user of our system finds itself in.

"Legal RuleML" is an extension to the Rule Markup Language (RuleML), in itself a markup language that was developed to express rules in XML for deduction, rewriting, and generally inferential reasoning tasks. The result is a canonical Web language for rules using XML markup and transformations between other rule standards/systems. To represent legal rules specifically, an extension to that language is developed, and its application demonstrated with a heterogeneous set of legal documents, ranging from legislation to contractual terms and condition statement. For copyright law, the chosen example is § 602 (b) of the US copyright law, Title 18, Chapter 6:

§ 602 (b) In a case where the making of the copies or phonorecords would have constituted an infringement of copyright if this title had been applicable, their importation is prohibited.

and its interaction with *§* 504 (c)(1) of the same law.

Legal RuleML then draws on insights from Legal Ontology research combined with semantic norm extraction based on Natural Language Processing (NLP) to assign these laws a formal representation that can be executed in a semantic web environment:

```
<Implies id="rule602b">
 <then>
     <prohibition>
           <Atom id="rule602b-prh1-atm1">
              <Rel>importation is prohibited</Rel>
                 <Var>z</Var>
           </Atom>
        </prohibition>
  </then>
  <if>
    <And>
       <Atom id="rule602-if-atm1">
             <Rel>copies or phonorecords</Rel>
             <Var>z</Var>
        </Atom>
        <Atom id="impl602-1-if-atm2">
             <Rel>without the authority of the owner of
                  copyright
</Rel>
          <Var>x</Var>
        </Atom>
      </if>
   </And>
 </Implies>
```

This representation renders the legal document into a machine readable format by laying bare its fundamental logical structure – the legal consequence "it is prohibited that…" can now be automatically detected by the computing environment, and it can be checked automatically if the conditions for "firing" the rule are give, and that legal consequence triggered. This is the section below the <if> sign, which lists the various preconditions that have to be met before the rule applies, and the logical relations between them.

As we have seen, after a long period of neglect computational representations of copyright law very recently garnered attention in the AI and Law community. The uptake by practice, in particular legal practice, remains however limited. There is one other approach to computational copyright though whose impact on practice and its commercial relevance can't be doubted, even though not all of the initial higher expectation shave been achieved, and in some sectors of the creative digital economy, their use has been in decline [see e.g. 26, 27, 28]. Digital Rights Management (DRM) "imitates" legal regulations by regulating access to digital resources. For Lessig, this new way to integrate legal compliance directly into software architecture became the archetypical example for an entire new form of Internet governance, regulation through software code. Despite this obvious connection between DRM and representation of legal concepts, interest in the AI and Law community has been muted. Only one paper, in the resources surveyed, attempts to leverage existing AI and law solutions for better DRM [28]. A small number of follow up studies citing this initial study exist, but are

published outside the core outlets for the AI and law community, such as [29, 30]. The reason for this limited response lies in the different design philosophies in traditional AI and Law research on the one hand, DRM on the other. AI and Law research tries to make the logical deep-structure of legal knowledge explicit, resulting ideally in isomorphic correspondence between the formal model and the natural language legal reasoning [31]. DRM by contrast only mimics the consequences of legal norms, but does not give an explicit account of how these results are reached based on a fully explicit and transparent reasoning process. This has considerable advantages for the development of commercial applications: Since legal knowledge is not explicitly represented there is no knowledge acquisition bottleneck, a persistent problem in expert system design. There are also advantages for scalability and efficient use of computational resources: fully explicit legal reasoning can make considerable demands in terms of runtime on the computing environment – having a system perform an explicit legal analysis, even if automated and supported by significant computing power, every time a digital photograph is opened or an e-book downloaded would mirror in the real world a situation where we ask for full legal advice when performing the most trivial of transactions. On the other hand, a recent decline in the use of DRM, and a focal point for widespread criticism of the technology is its frequent overreach [34]. It prevents transactions and operations the buyer of a digital object would be legally entitled to perform. This is the dark side of the same coin – absent an explicit and fully formulated representation of the legal environment, "dumb" DRM can catch only a small aspect of the legal meaning of "having a license" or "buying a file". DRM's relation to copyright is similar to that of a traditional lock to property law – a very rough approximation, but in the same way in which a physical lock will prevent entry also for people who are entitled to, so does DRM often prevent legitimate uses.

We face as a result a dilemma: DRM is "dumb" but has a track record of application in practice, even though one marred by constant criticism. Traditional legal AI is intelligent, and in principle capable of addressing the problem of "overreach" in DRM, but struggles to develop application of commercial strength. Semantic web approaches are an important development to bring the two traditions of research together, and out own study also follows this approach – can we learn from the relative failure of legal AI to reach legal practice, and the success of DRM, while maintaining the commitment to fully explicit representations of legal knowledge –something tah ultimately should also benefit attempts to automatize the management of digital rights.
while extending the expressive power of the hypertext markup languages to cover legal knowledge is a highly desirable goal in itself, research such as Palmirani's do not normally describe a specific business case in legal practice. Finding relevant legal information on the net, a legal information retrieval task, and exporting that information into a variety of formats, is an obvious application. This too however remains highly generic, with the same type of task experienced by judges, solicitors or laypeople. While the generic, context independent use of the formalism therefore makes it theoretically suitable for a particularly broad and heterogeneous range of legal tasks, it needs for its adaptation a rather demanding set of skills on the user side. Law firms or other potential users need to develop an understanding of the benefits and potential of ontology driven semantic web technology; they need to understand the particular advantages that the proposed formalism offers them; they need to identify within their institution those knowledge-intensive practices that would benefit from an automated or semiautomatic approach; and finally they need to find a way to "match" there internal practices.

This initial discussion leads us to the formulation of a research strategy that in some ways shares the aim of DRM – lowering the costs of copyright enforcement by automatizing some of its operation, but which remains in the methodological tradition of legal AI research and its commitment to fully explicit representations of legal knowledge. To respond to the problem of scalability mentioned above, this means we have to reorient the approach away from the point of purchase or use of a digital resource – where time constants make a full fledged legal analysis impossible, and back to an environment where, in relative terms, time does not matter that much – the law firm.[3] By optimising the process of licensing and litigation preparation through judicious use of automation, costs to the legal system, and also to clients and customers, should fall. However, generic solutions that do not represent specific contexts of legal transactions, as those discussed above, have struggled to find the interest of legal practice. A different methodology I therefore needed, one that aligns the development of a formalism more closely with the actual practices and constraints of an actual law firm. One of the reasons for the failure of legal AI to reach legal practice, so we will argue, was a problematic focus on "judicial" reasoning – legal AI models typically judges and their reasoning, not lawyers assisting their clients in non-adversarial settings. Just like DRM, we will therefore focus initially not on adjudicative situations where a conflict already exists, but at the earlier stage of drafting and managing licenses. Unlike DRM though, we focus on more complex licensing arrangements that require direct, real-time interaction between lawyers and clients.

However, to establish the feasibility of our approach, modelling copyright licensing directly proved overly challenging. We will argue that to model licensing and contract formation at all, the conceptual vocabulary of ontology driven representations of legal knowledge have not just to be increased, an entire shift of focus is required. The substantial degrees of freedom that copyright licenses allow (e.g. complex arrangements on moral rights) mean that they are less suitable to develop an "ideal type" of formal models of legal transactions, something that we will argue is a necessary extension of the conceptual vocabulary of legal ontologies. In a somewhat circuitous way, we will therefore discuss the empirical results of a study in a much less ocmpex form of transaction first, and develop on this basis a ne formal model of a "legal transaction". We ten indicate how this generic model already addresses one of the main obstacles that stood in the way of law firm's use of legal AI, and indicate how this "bare bones" concept can then be extended to cover also the more complex and demanding field of copyright licensing.

In summary, to respond to the challenge identified by Oskamp and Laurits, Orlando Conetta carried out a case study within a magic circle[4] law firm headquartered in the City of London, with the goal to better understand the legal, technological and sociological reasons for the lack of uptake of legal AI in commercial practice. Based on this study, we develop a new concept, Transaction Configuration as an abstraction suitable for AI modeling that captures the main task commercial lawyers face in the performance of their work.

---

[3] of course, time matters also in law firms. But the demands on time when discussion scalability of computational resources is measured in milliseconds, well below what is noticeable by individual humans beings and of relevance mainly when million of operations need to be performed simultaneously.

[4] The magic circle refers to the leading commercial City law firms, who are: Allen & Overy, Clifford Chance, Freshfields Bruckhaus Deringer, Linklaters, and Slaughter and May.

The main hypothesis of the study was derived from André Valente's influential book [14]. Through the explicit adoption of the CommonKADS knowledge engineering methodology, Valente proposes an active role for knowledge engineering in the exposition of legal tasks. Knowledge Acquisition and Documentation Structuring (KADS) is a structured way of developing knowledge-based systems that was developed at the University of Amsterdam, and made an important contribution to address the relatively unsuccessful reception of expert systems in business practice. It remains to this day one of the most important methodologies to provide support for the production of Knowledge Based Systems (KBS) in an industrial approach. It comprises a set of models which guide the production of the system in a systematic and structured way, which assists in particular the tasks of analysis and the transformation of expert knowledge into a form that can be read and executed by machines. From this starting point, Valente develops a new template model for legal normative assessment. In line with this approach, the case study was designed to use Valente's engineering methods to apply legal normative assessment using LKIF, the Legal Knowledge Interchange Format [31, 32]. We briefly discussed above the role of the semantic web to build usable, computational representations of law. LKIF is part of this effort. Developed by the European ESTRELLA project, it takes the Web Ontology Language (OWL) for representing concepts as a starting point, and enriches it with XML schemata for rules and arguments. It was designed with the goal of becoming a standard for representing and interchanging policy, legislation and cases, including their justificatory arguments, in the legal domain.

Section three describes the CommonKADS feasibility study of Eurobond transactions. By applying LKIF to debt selling regulations, it was found that the exhaustive modelling of primary and secondary norms, and legal normative assessment, is of little use to the commercial lawyer during the negotiation of a contract, when future facts are unknowable.

In light of this, section four presents *Transaction Configuration* (TC) as a new legal task type, and describes how it can be applied against a restricted form of LKIF. It will be shown that TC provides the relevant context for normative legal assessment in commercial practice by determining best practice from the subjective 'features' of a transaction. As such, it will be argued that well engineered *Transaction Configuration* systems could be cost-effectively developed with the "immediate benefits" to lawyers that Oskamp and Laurits proscribe. We then outline an extension of our study currently carried out within the CREATE network[5], focussing on copyright licensing as the transaction under investigation. This new application offers several advantages: the digital economy relies heavily on copyright law, which makes it particularly natural to think of computer assisted technology to manage complex transactions, as the object of the transaction can in theory be *used* in the transaction directly, not just *mentioned* though a description. This enables novel ways to assist in "lawful service engineering" and reduce transaction costs.

Second, initial explorative talks with in-house copyright lawyers in the music industry identified similar themes to the ones we to the way in which Eurobond transactions are managed in commercial law firms. Here too, costs are high, as is complexity of the task. Nonetheless, initial explorations also indicated substantial differences between complex copyright licensing and the Eurobond deals described in this paper. In particular, the role of potential litigation with third parties is much greater

[5] www.create.ac.uk

in copyright licensing, which will therefore be a real test to our contention that transaction configuration, rather than argument modeling, is the core need of commercial lawyers that legal AI should and can address.

Another main difference that emerged is the role of explicit argumentation and negotiation. In Eurobond transactions, the clients are professionals with very clear ides of the goals they want to achieve, and goal conflict is rare. In copyright licensing by contrast, the ultimate client, artists, often gain an explicit understanding of their goals only in the process of deliberation with their lawyers, with conflicts between their commercial interests and their artistic vision, protected through "moral rights", frequent. We are currently working within CREATE on an interview based study of artist's goals and motivation, and the way in which they utilize copyright to achieve these goals. This will complement a CommonKADS based study of the practice of copyright lawyers. We predict that unlike in Eurobond Transactions, argumentation will play and important and dual function in this domain. We expect to find clear instances of collaborative deliberation dialogues between clients and lawyers, involving goals and values. Secondly, in preparation of possible litigation persuasion dialogues are likely to take place in parallel. Under these conditions, it will be natural to extend the approach presented here by the explicit argumentation based approach of the CARNEADES model [16,17]. In the Eurobond domain by contrast, the added value of an argumentation based approach seemed much less clear, and our empirical study found fewer instances of argumentation, which explains why, we do not use the CARNEADES formalism here, to emphasize the difference between a court focused, litigation centered approach to AI where argumentation is key, and our "service oriented" model where its role is less obvious.

2. Case Study Design

Estrella's[6] Legal Knowledge Interchange Format (LKIF) is a layered ontology, written in OWL, which has abstracted the terminological knowledge necessary to articulate legal arguments using situational, mereological and epistemological frameworks [2]. The project team also developed Harness [4, 15], a Protegé add-on to allow users to run simulations using a compatible OWL-DL reasoner. For these reasons it was decided that LKIF would be the most appropriate toolkit to apply, and was consequently set as the study's unit of analysis. During the course of the CommonKADS feasibility study various participants were interviewed over several days in order to complete the respective OM-1 and OM-2 sheets. A sanitised list of those participants is listed below.

- Partner A: Experienced banking partner, champion of legal knowledge systems.
- PSL A: Non fee-earning lawyer responsible for the drafting of document assembly templates.
- Strategy Manager A: Member of the strategy team, which reports to the senior management group of the firm.
- Knowledge Manager A: Senior member of the cross-function knowledge management department.

[6] European Commission Project: IST-2004-027655 http://www.estrellaproject.org/

– IT Consultant A: Member of IT team responsible for liaising with functional business the development of new IT solutions.

3. LKIF and Selling Restrictions

The chosen transaction type for analysis was a Eurobond Medium Term Note (EMTN) Programme Establishment. In this transaction a company wishes to issue debt as bonds, but also wants to retain a degree of flexibility as to when and how tranches of debt are issued. The programme establishment sets up all the relevant master documents and authorisations, so facilitating the rapid issue (or drawdown) of new bond tranches over the lifetime of the programme [5 p. 191].

3.1. Process and Roles

The programme establishment is usually followed by the first issue drawdown, with its own process. In the case of the programme establishment the goal is to negotiate and sign the relevant agreements and master documents. While there are procedural elements, the overall process is more akin to a dynamic schedule that fits around the negotiating parties and works towards a signing date. An observation that was indirectly matched by strategy manager A in relation to earlier attempts by the firm to map transactions as workflows. Then, as in this instance, it was found that transactions have three main stages; design, negotiate and draft, then complete. However, "the devil is in the detail" and at each new level, and for each new task "there's all the loops", so manager A that would be conditional on other sets of factors. It was the belief of strategy manager A, that these "loops" took the problem beyond the capabilities of traditional workflow theories and tools.

The two documents considered in this study were the base prospectus and the programme agreement; where the latter is the main agreement that will bind the issuer, arranger and dealers for the duration of the programme. A prospectus is a disclosure document that will be available to all potential lenders during a drawdown, and must conform to strict standards set in law. If the issue is to be listed then the prospectus will also need to adhere to the requirements of the respective exchange. In this instance, during the programme establishment the base prospectus will be submitted to the relevant listing authority for pre-approval, so as to facilitate the rapid listing of subsequent drawdowns.

Before examining further the types of tasks lawyers perform in the course of a transaction, the initial challenge was to determine the relevance of legal normative assessment, as devised by Valente, and implement in Harness an aspect of the EMTN Programme. A small subset of the law relating to the marketing of debt was examined, and compared against the controls used in knowledge assets to guide lawyers.

The Financial Services and Markets Act 2000 (FSMA) imposes restrictions on the sale of debt in the UK. In 2005 FSMA was amended by the Prospectus Regulations 2005 in order to implement Directive 2003/71/EC (Prospectus Directive). Section 85(a) of FSMA is the ratified form of Article 3.1, of the Prospectus Directive, that provides that it is unlawful for a transferable security to be offered to the public unless an approved prospectus has been made available before the offer. Section 86 provides a set of exceptions to this prohibition to account for offers clearly intended for professional investors, and not just the general public. Meanwhile, s19 of FMSA

imposes a “general prohibition” on regulated activities, which are defined in Article 5(1) Regulated Activities Order 2001 (RAO) as “accepting deposits” to be repaid after it has been used to finance the activities of the borrower. The question as to whether this relates to debt issues, to the public or otherwise, becomes more involved, as Article 9(1) deems money received in exchange for debt securities as not to be a deposit for the purposes of Article 5(1). However, Article 9(2) will still apply to considerations received for the issue of commercial paper, which denote debt securities with a maturity of less than one year.

In applying Harness to this task, while remaining mindful of individuation [13 pp120-128], a secondary goal was to assess how the Harness framework could translate into an LKIF representation of a legal transaction. As per the normative assessment functional specification, generic cases were established to represent the circumstances against which each individuated norm would apply.

Focusing on FSMA, a generic restriction class denoting any debt issue, forbidden by s85(1), was established by setting it equivalent to the new concept of a debt issue. Then a more specific debt issue was defined with an extra restriction to have published a prospectus document. Both classes are represented below:

```
GC85(1) ≡ Debt_Issue
GC85(1)P ≡ Debt_Issue and published some Prospectus
```

These cases could then be associated to subclasses of the norm concept in Harness to set GC85(1) as forbidden and GC85(1)P as allowed, respectively. The requirement to publish a prospectus was represented as an obligation, with it permissive element extracted away from the prohibited context.

As the exercise progressed the restrictions became more involved. For example, section 86(1)(b) releases an issuer from the requirement to publish a prospectus, where such an issue is directed at fewer than one hundred qualified persons. However, when reviewing the respective provision in the clause template, there was an additional contractual obligation against the issuer to obtain the prior consent of all the dealers for the offer. The class restriction below attempted to represent this mix between the contractual and statutory obligations.

```
GC85(1)(b) ≡ Debt_Issue and dealer some (Agent and author value
market_to_nonqualified_investor_consent) and offered_to max 100
NonQualifiedPerson
QualifiedPerson ≡ (Legal_Person and feature min 2 Feature) or
(Authorised_Person or Professional_Investor or Regulated_Person)
```

A non-qualified person (NonQualifiedPerson) was defined as disjoint to the qualified person (QualifiedPerson) restriction. However, company features such as a balance more than 45m euro could only be represented properly as A-Box individuals, and so the exercise identified the need to represent meta-knowledge in a framework.

Moreover, as concepts are inferred within a hierarchy, or an acyclic graph, it would prove impossible to restrict OWL classes based upon specific individual’s reference to each other [8], and represent transactions between dealers and issuers. Although, with respect to the hybrid design pattern in Harness, a representation could go beyond OWL-DL and use rules to solve this problem, but at the cost of further technical complexity. With regard to normative assessment, modelling 85(1)(b) showed how Generic Cases would mix scenarios regulating actions that have yet to take place (e.g. offer to fewer that 100 investors), with immediate obligations that can manage future risks (e.g. seek dealer consent). Where the purpose of the system was to produce a normative model for a judge, or litigator, deciding how parties should have behaved

in the event of a breach, then this combination of contractual and legislative norms would be appropriate. Although, in light of the context of use, such knowledge would and could not be applied to events that have yet to take place, before the transaction is complete.

4. Transaction Configuration

The first phase of the study had shown that the main task of a transactional lawyer is not to construct arguments regarding past events, but instead to apply best practice in the construction of new agreements to regulate future actions. While normative assessment could not adequately describe the main legal task pertinent to a commercial negotiation, there was evidence that it was relevant as a part of a wider task specification, such as when specifying the appropriate consents required.

CommonKADS sheets OM-3 and OM-4 are designed to focus the knowledge engineer on the most specific aspects of the task for modelling. However, examining a transaction at such a level of detail would have run the risk of encountering strategy manger A's "loops" and not elucidate a more general description of an activity that would be applicable against other transactions. It was found that by considering abstracted classes of knowledge assets however, a pattern of performance emerged that partially matched an existing CommonKADS task template: the configuration "propose-and-revise" method offered by Schreiber et al[12 p.149].

For example, in the case of a checklist, the lawyer would first have to select the appropriate template, apply the subjective details of the transaction against it, and then assess any legal or commercial risks identified. Whereupon the lawyer could either refer such risks back to the client, or seek to better understand the legal or commercial basis of the checklist item from which the risk was flagged. PSL A, who was responsible for the design and upkeep of a set of automated document assembly templates, stressed the complexity involved in maintaining even a single document assembly template.

> "we have broad key questions which are the kind of key structural questions, so is it [the loan] syndicated or bilateral? How many facilities are you doing? What types? Is it English law or not? I mean, thinking of the statistics of it, [we] are already producing a large number of combinations."

Furthermore, when discussing the challenges of indexing the artefacts from completed transactions, IT Consultant A explained the nature of "feature" information.

> "So the feature is really around the dissection of the matter and the key terms that they're interested in, so that plays a big part in this, and obviously party information, so who was involved in the deal [...] because you might say, who have we worked for in Saudi Arabia in the last six months?"

These dimensions were similar to those encountered when manually filtering a checklist to determine the appropriate actions. Many "features" were free text, such as party names, while others were categories selected from fixed sets, like in the case of the jurisdiction. The checklist could then be filtered on the basis of these set choices, and then be applied across various types of transaction that have the same attributes in common, much in the same way property-centric ontologies in OWL infer memberships to classes. This observation led to three conclusions.

(1) *Transaction types are dynamic*: IT Consultant A explained the operational nature of transaction types in the firm.

> "Transaction types actually do exist [...] and from my understanding the driver

for that is to provide the management information"

Although, from the functional side of the business (i.e. the lawyers), categories of transactions described with features are less absolute, and a single instance of a deal could be described against multiple transaction types.

(2) *Transactional lawyers already use knowledge engineering*: Following from above, practitioners consider transactions by their parts, rather than as discrete products, and develop knowledge with respect to these parts. Knowledge is primed to be reused within the context of deals classified against a range of transaction types, and not just within a single packaged service.

(3) *Current engineering methods are inefficient*: Due to the implicit nature of engineering within legal practice, current efforts lack transparency and are not guided by theories or methodologies. With the result that knowledge encoding is duplicated across different knowledge types. For example, a category to describe if a loan is syndicated or otherwise could be set by a PSL drafting an automated loan document, as well as a partner devising a new checklist template, and separately again by the knowledge officer when tailoring a metadata schema for loan transactions. Consequently, during a deal, the lawyer will also be required to apply the same deal information separately into each knowledge source.

In light of this duplication of engineering effort, it was concluded that the value that could be delivered by a unified knowledge framework, pursuant to particular legal activity, would be far greater than the sum of the individual knowledge assets maintained currently. To that end the study developed a framework to describe transactions with LKIF concepts, presented below as the *Transaction Framework*.

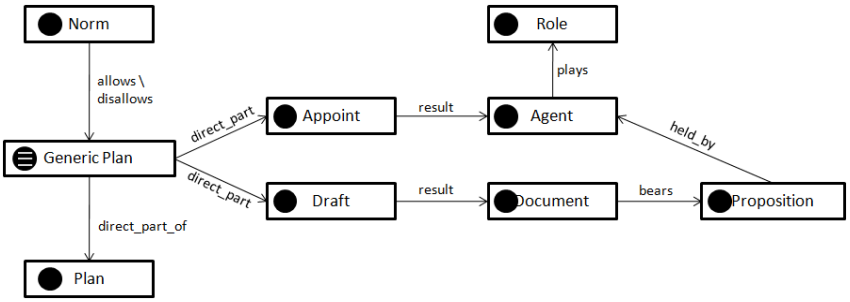

Figure 1. LKIF Transaction Framework

As most features encountered during the study were factual in nature, Propositions were isolated from propositional attitudes, relevant to describing arguments, and allocated the function of representing transaction features such as those encountered above. A Proposition value relationship can represent normal literals, enumerated sets, and be empty or represent a default organisation position. Moreover, a Proposition could be set as equivalent to Temporal Occurrence within LKIF, so facilitating an ordering of features representing checks. All Propositions must either be linked to a source Document, or an Agent, or both. Despite its title, a Document can represent any relevant assertions and does not need to be part of a functional or legally binding document. For example a Document would include guidance notes, clauses, checklists, or any other content. Within a transaction a Document or Agent is the result of an LKIF Action to respectively Draft or Appoint.

The mereological layer in LKIF is used to represent the groups of Actions deemed to be relevant to a part of a transaction, or a Plan. Thus a transaction is represented as

Appoint and Draft Actions that are parts of Plans (as transaction parts) that, in turn, make up other Plans. Mirroring the dynamic classification of transactions by lawyers, transaction parts can be inferred through OWL restriction classes, which are set as equivalent to a Plan and can be deemed to be allowed or disallowed by Norms. This structure allows for flexibility, where the engineer could model a Generic Plan, set as disallowed by a procedural Norm, and link it to Actions that would instruct the lawyer as to appropriate countermeasures. In other words, Norms need not all be based upon laws, but where required the Framework still provides the concepts necessary for legal normative assessment.

An engineer encoding a Plan would do so with RDF individuals denoting specific Agents, Documents and Propositions, linked to an individual nominated by an OWL restriction class. A Plan is then deemed *relevant to* a transaction when it is successfully *inferred against* an existing restriction class. However, a *Transaction Framework* is clearly an example of a meta-model structure (that OWL-DL rejects for decidability), as it is described with both OWL classes and RDF individuals in the T-BOX. Any implementation would consequently encounter a technical challenge to determine the best means to properly infer an A-BOX instance typed against a particular *Transaction Framework*. Punning would be one approach [10]. The *Transaction Framework* structure raises however the current shortcoming of OWL DL to explicitly handle meta-models. The *Transaction Framework* was designed in the context of the common configuration activity that was observed in the use of checklists, styles, and document assembly templates. As stated above, this configuration pattern was also found to be similar to a variation of the propose-and-revise (PaR) CommonKADS activity template. Presented below is Transaction Configuration, a new activity template, based upon PaR, that offers a new perspective on the process of conducting a legal transaction, and places Valente's normative assessment task within a practical context.

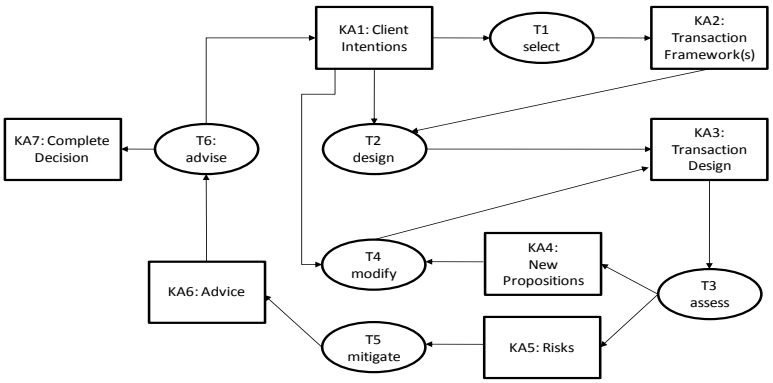

Figure 2. Function Structure for Transaction Configuration

Figure 2 is a directed graph representing the CommonKADS function structure for *TC*, and illustrates how knowledge flows between the decomposed tasks of the *TC* function. The corresponding control structure is described in detail below and assumes the domain schema to be the *Transaction Framework* specified in section 3.1 above. All references to *Transaction Framework* entities are capitalised.

```
while KA7:Complete Decision = false do
   T1:select (KA1:Client Intentions -> KA2:Transaction Frameworks);
```

```
    T2:design (KA1:Client Intentions + KA2:Transaction Frameworks ->
KA3:Transaction Design);
    T3:assess (KA3:Transaction Design -> KA4:New Propositions +
KA5:Risks)
while KA4:New Propositions do
       T4:modify (KA4:New Propositions + KA1:Client Intentions ->
   KA3:Transaction Design);
       T3:assess (KA3:Transaction-Design -> KA4:New Propositions +
   KA5:Risks)
end while
    T5:mitigate (KA5:Risks -> KA6: Advice)
    T6:advise (KA6:Advice -> KA1:Client Intentions + KA7:Complete
Decision)
 end while
```

T1: Select Transaction Frameworks

A transaction is instigated upon the client recognising a need and describing intentions to the lawyer. In practice these intentions could be described within a standard form, or term sheet, and are denoted in the activity as knowledge asset 1(KA1). On the basis of these intentions the lawyer should be able to select the top level *Transaction Framework*s deemed relevant (KA2). This list is reviewed after each cycle (see T6 below), and frameworks can be removed the list if intentions change. For example, in the case of an EMTN programme with a subsequent drawdown, the lawyer could select a EMTN Programme Establishment Framework, along with a EMTN Drawdown Framework for each subsequent transfer. The ease with which a lawyer finds appropriate *Transaction Framework*s will depend on the skill of the legal knowledge engineer to know the optimal level at which a Plans should be ring-fenced.

T2: Design Transaction

Based upon the client intentions (KA1), the lawyer defines values for the Propositions in the transaction design (KA3), which have been applied onto the design by the selected *Transaction Framework*s (KA2). If a framework has been removed from the list (KA2), then the respective graph is also removed from the design (KA3).

T3: Assessment

Upon defining a Proposition in the design (KA3), the system will automatically infer all applicable Generic Plans within the *Transaction Framework*s, which then determine the Documents that are deemed relevant to the circumstances of the deal. In doing so, the user may be presented with more Propositions (KA4), which belong to the Documents that were previously hidden from the Generic Plan. Similarly, the task will generate a risk report (KA5) by performing normative assessment against those Generic Plans that are associated to a normative status (e.g. allowed or disallowed). However, instead of inferring an overall assessment deeming the transaction either allowed or not, the task outputs a detailed explanation of the relevant risks, and offers best practice and guidance to the lawyer.

T4: Modify Transaction Design:

Following assessment (T3), there may be new Propositions (KA4) for the lawyer to review, the values to which are set within the transaction design (KA3) before another assessment (T3). This assessment-modify cycle shall continue until the user is satisfied that as many Propositions as possible have been defined.

T5: Mitigate Risks

Based upon the explanations and guidance presented in the risk assessment (KA5), the lawyer shall compose advice (KA6) to take to the client. This may include suggested

changes to the client strategy, or intentions (KA1), or be more basic advice notifying the client of the status of ongoing actions.

T6: Advise Client

Based upon the advice (KA6) composed, the lawyer may agree with the client to either complete the deal, or alter the current strategy (KA1) and continue on to another cycle. Alternatively, it may be that certain propositions, possibly relating to checklists, are still pending. In which case the cycle shall repeat, with the complete decision flag (KA7) set as false, until the respective Propositions are defined (T4).

5. Conclusion and future work: From Eurobonds to Copyright

LKIF proved to be competent at describing the rich features of a highly complex legal transaction, and is an ideal platform from which to test and prototype legal frameworks. *TC* allowed us to shift the emphasis away from the lawyer as judge or litigator whose main job is to argue about past events, to what the empirical part of the study identified as the most common role of a commercial lawyer: shaping a client's actions to help him achieve future goals. The practical relevance of the wider *TC* task template though can only be established through further empirical development of domain applications. For instance, the *TC* template presents the requirement to model and reason with client intentions (see KA1). However the *Transaction Framework* does not make use of propositional attitudes, a material element of LKIF rules (see Carneades [7]), which could conceivable be brought to bear on the challenge KA1 presents. This limitation is further apparent when it is considered how *TC* may be applicable beyond the execution of commercial transactions, in the preparation of litigation disputes that could arise.

To address these questions, we are currently expanding the model to drafting of copyright license that involves different types of protected works such as music, words and performances. The framework will account for the perspective of a lawyer assisting in the completion of a new contract, as well as that of a lawyer defending or asserting claims in a subsequent dispute arising from the same contract. By taking a 360 degree perspective on the same legal domain, it is anticipated that the respective strengths and weaknesses of the *TC* task template and *Transaction Framework* can be examined at a higher level of resolution. Extending the approach to IP licensing will also address one of the possible objections to the case study chosen in our paper here. Eurobond Medium Term Notes operate in a highly complex legal environment, and in the "shadow" of litigation when things go wrong. Nonetheless, the actual danger of litigation is low, reducing the need to anticipate possible legal arguments in an adversarial setting. IP licensing, while also "cooperative" in the sense that several parties try to maximize their benefits through contracts, is much more likely to result in litigation with third parties not involved in the actual drafting. Unlike in the present example, the interests of the clients will also be much more complex, with financial rewards being only amongst several goals they may have in addition to more intangible objectives such as fame, the integrity of their art or opinions about their use for political, commercial or other problematic purposes. This creates a much greater variety of possible configurations, and makes their reduction to checklists a much more problematic exercise. The *Transaction Framework* is hence only a starting point for any practical implementation, and it would be expected and appropriate for it to develop beyond OWL and LKIF in order to build reliable and scalable *TC* systems.

6. REFERENCES

[1] Apistola, M., & Lodder, A. R. 2005. Law firms and IT. Towards optimal knowledge management. *Journal of Information, Law and Technology*, *2*(3), 1–28.

[2] Boer, A. 2009. LKIF Core: Principled Ontology Development for the Legal Domain. *Law, Ontologies and the Semantic Web: Channelling the Legal Information Flood*, 21.

[3] Breuker, J., & Hoekstra, R. 2004. Epistemology and ontology in core ontologies: FOLaw and LRI-Core, two core ontologies for law. In *Proceedings of EKAW Workshop on Core ontologies. CEUR*.

[4] Breuker, J., & van de Ven, S. 2008. Deliverable 4.6, Developing HARNESS, Towards a hybrid architecture for LKIF. DOI = http://www.estrellaproject.org/doc/Estrella-D4.6.pdf

[5] Choudhry, M. 2006. *An Introduction to Bond Markets* (3rd ed.). John Wiley & Sons.

[6] Förhécz, A., & Strausz, G. 2009. Legal Assessment Using Conjunctive Queries.

[7] Gordon, T. F., and Walton. D. 2006. The Carneades argumentation framework. *FAIA* 144 195.

[8] Hoekstra, R., & Breuker, J. 2008. Polishing diamonds in OWL 2. In *Proc. of the 16th Int. Conference on Knowledge Engineering and Knowledge Management* (EKAW 2008), LNAI/LNCS. Springer Verlag.

[9] Horridge, M., Drummond, N., Goodwin, J., Rector, A., Stevens, R., & Wang, H. H. 2006. The Manchester owl syntax. *OWL: Experiences and Directions*, 10–11.

[10] Jekjantuk, N., Groner, G., Pan, J. Z., Zhao, Y. 2010. Modelling and validating multilevel models with owl fa. In: Proc. of the 6th International Workshop on Semantic Web Enabled Software Engineering

[11] Oskamp, A., & Lauritsen, M. 2002. AI in Law Practice? So far, not much. *AI and Law, 10*(4), 227-236.

[12] Schreiber, G. 2000. *Knowledge engineering and management*. Cambridge Mass.: MIT Press.

[13] Susskind, R. 1987. *Expert systems in law*. Oxford University Press, Inc.

[14] Valente, A. 1995. Legal knowledge engineering: A modelling approach. Penn State Press.

[15] van de Ven, S., Hoekstra, R., Breuker, J., Wortel, L., & El-Ali, A. 2008. Judging Amy: Automated legal assessment using OWL 2. *Proceedings of OWL: Experiences and directions (owled 2008 eu), 44*.

[16] Gordon, T. F. (2010). An Overview of the Carneades Argumentation Support System, Dialectics, Dialogue and Argumentation, ed. C. Reed and C. W. Tindale, London: College Publications, 145-156

[17] Atkinson, K. and Bench-Capon, T. (2007). Practical Reasoning as Presumptive Argumentation Using Action Based Alternating Transition Systems. Artificial Intelligence 171, 855–874

[18] Gordon, T., (2011) Analyzing open source license compatibility issues with Carneades, Proceeedings of the 2011 ICAIL Conference ACM 2011 p 51-55

[19] Governatori, G., Rotolo, A., Villata, S., & Gandon, F. (2013). One License to Compose Them All. In The Semantic Web–ISWC 2013 (pp. 151-166). Springer Berlin Heidelberg.

[20] Rotolo, A., Villata, S., & Gandon, F. (2013, June). A deontic logic semantics for licenses composition in the web of data. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law (pp. 111-120). ACM.

[21] Bavota, G, et al. "The MARKet for Open Source: An intelligent virtual open source marketplace." Software Maintenance, Reengineering and Reverse Engineering (CSMR-WCRE), 2014 Software Evolution Week-IEEE Conference on. IEEE, 2014

[22] Contissa, G., & Laukyte, M. (2008). Legal Knowledge Representation: A Twofold Experience in the Domain of Intellectual Property Law. In *JURIX* (pp. 98-107)

[23] Palmirani, Monica, et al. (2011) Legalruleml: Xml-based rules and norms. Rule-Based Modeling and Computing on the Semantic Web. Springer Berlin Heidelberg, 298-312

[24] Bourcier, D, Casanovas P, Dulong de Rosnay, M, Maracke C (2010) (Eds.) Intelligent Multimedia: Managing Creative Works in a Digital World, European Press Academic Pubishing

[25] Rosenblatt, W, Mooney, S and Trippe, W, (2001). Digital rights management: business and technology. John Wiley & Sons, Inc., 2001

[26] Becker, E. (Ed.). (2003). *Digital rights management: technological, economic, legal and political aspects* (Vol. 2770). Springer

[27] Ashtar, Reuven. "Licensing as Digital Rights Management, from the Advent of the Web to the iPad." *Yale JL & Tech.* 13 (2010): 141

[28] García, R, Gil, R and Delgado J (2007). "A web ontologies framework for digital rights management." Artificial Intelligence and Law 15.2 137-154.

[29] Prandoni, C., Valentini, M., & Doerr, M. (2009). Formalising a model for digital rights clearance. In *Research and Advanced Technology for Digital Libraries* (pp. 327-338). Springer Berlin Heidelberg.

[30] Zhang, X. M., Liu, Q., & Wang, H. Q. (2012). Ontologies for intellectual property rights protection. *Expert Systems with Applications*, *39*(1), 1388-1400.

[31] Alexander, B. O. E. R. (2009). LKIF Core: Principled ontology development for the legal domain. Law, ontologies and the semantic web: channelling the legal information flood, 188, 21.

[32] Hoekstra, R., Breuker, J., Di Bello, M., & Boer, A. (2007). The LKIF Core Ontology of Basic Legal

Concepts. LOAIT, 321, 43-63.
[33] Bailey Jr, Charles W. "Strong copyright+ DRM+ weak net neutrality= digital dystopia?." *Information Technology and Libraries* 25.3 (2013): 116-127.
[34] Favale, M. (2011). Approximation and DRM: can digital locks respect copyright exceptions? International Journal of Law and Information Technology, 19(4), 306-323.
[35] Akester, P. (2009). *Technological Accomodation of Conflicts Between Freedom of Expression and DRM: The First Empirical Assessment*. Centre for Intellectual and Information Law, University of Cambridge.

RCUK Centre for Copyright and
New Business Models in the
Creative Economy

College of Social Sciences / School of Law
University of Glasgow
10 The Square
Glasgow G12 8QQ
Web: www.create.ac.uk