Controlled Natural Language rendering of Copyright Ontology licenses

Roberto García and Rosa Gil

Universitat de Lleida, Spain
{rgarcia,rgil}@diei.udl.cat

Abstract. Digital Rights Management and, more concretely, Rights Expression Languages can benefit from formal semantics based on ontologies. However, this increased level of formality can make licenses even harder to read than those based on less formal solutions like XML Schemas. In order to overcome this usability issue, the Copyright Ontology is based on an ontology-engineering pattern that follows a Natural Language-Oriented approach. Licenses based on this ontology are based on verbs connected to agents, objects, time points, places, etc. This sentence-like structure makes it possible to render them as a Controlled Natural Language (CNL). This rendering has proved to be more appealing to end-users and the expressions built using the proposed pattern more usable for them.

1 Introduction

Most Rights Expression Languages (REL) for Digital Rights Management (DRM) express content rights in a machine-readable form, commonly using XML technologies. Consequently, they are usually based on just a formal syntax but not formal semantics, which makes them not as machine-readable as it might be possible using formal semantics approaches, e.g. those based on ontologies[1, 2].

The main advantages of the presence of formal semantics are that it facilitates interoperability and the development and maintenance of systems capable of working with those semantic expressions. Interoperability is facilitated because it is possible to work at the conceptual level, where different kinds of relations can be established and make machines reason about them. It is also possible to facilitate the implementation and maintenance of DRM systems because it is easier to go from the specification to the implementation, i.e. the application can be built on top of the semantic descriptions of the domain and profit from reasoning mechanisms so they are in charge of part of the "implementation" effort.

However, the use of formal languages makes it more complicated for those not familiarised with them. When talking about developers, this is just a matter of training. However, when talking about end-users, it is a usability problem that must be avoided. The same happens with RELs based on XML. End-users should not face the XML expression that models the license. It is possible to use some sort of template that transforms the license into more palatable renderings.

However, these transformations get quite complex and difficult to maintain due to RELs complexity.

In fact, most users prefer DRM licenses in Natural Language form like the ones proposed by the Creative Commons (CC) initiative [3]. Despite the CC approach is not formal and is constrained to a set of predefined licenses, it is currently the more extensively used DRM system nowadays. The preferred and mandatory form of CC licenses is Natural Language in simplified or full-legal form, while there is a simple formal version intended for archiving purposes.

Therefore, there is a trade-off between formal DRM languages, which facilitate automating DRM systems but are less usable for end-users, and informal DRM languages, which are preferred by end-users due to their greater usability but make automation a very complicated issue.

The Copyright Ontology[1] [4] tries to challenge this trade-off proposing a quite formal approach based on ontologies but at the same time applies a Controlled Natural Language [5] approach that facilitates rendering expressions based on this ontology in a quite readable form. This approach is based on a pattern for ontology engineering that organises ontology entities in a similar way to those we found in Natural Language sentences.

This is a generic pattern but, in our case, it produces a formalisation of the copyright domain for automated DRM and provides some guidance for ontology creation. On the other hand, it facilitates generating a Controlled Natural Language rendering that, while keeping its formal nature, makes licenses easier for end-users. This pattern is described in Section 2, while in Section 3 it is shown how it has been applied to build the Copyright Ontology. Then, Section 4 presents the transformation from Copyright Ontology to CNL expressions and some examples from the copyright domain. Finally, the conclusions and future work are presented in Section 5.

1.1 Related Work

The DRM Watch review on DRM standards [6] shows that interoperability is a key issue for DRM systems. For instance, it arises in the content distribution scenario when a users want to consume content in any of the devices they own. Interoperability is also critical in the organisation scenario, when content flows through organisations or external content is used in order to derive new one.

The main response to DRM interoperability requirements has been the settlement of many standardisation efforts. The main ones are ISO/IEC MPEG-21[7] and ODRL [8], and in both cases the main interoperability facilitation component is a Rights Expression Language (REL).

The REL is a XML Schema that defines the grammar of a license modelling language, so it is based on a syntax formalisation approach. There is also the MPEG-21 Rights Data Dictionary and a ODRL Data Dictionary Schema (DD) that captures the semantics of the terms employed in the REL, but it does so without defining formal semantics [9].

[1] Copyright Ontology, http://rhizomik.net/ontologies/copyrightonto

This syntax-based approach is also common to other DRM interoperability efforts and one of main causes of the proliferation of interoperability initiatives that cannot interoperate among them, like in the e-books domain [10]. Despite the great efforts in place, the complexity of the copyright domain makes it very difficult to produce and maintain implementations based on this approach.

The implementers must build them from specifications that just formalise the grammar of the language and force the interpretation and manual implementation of the underlying semantics. This has been feasible for less complex domains, for instance when implementing a MPEG-4 player from the corresponding specification. However, this is hardly affordable for a more complex and open domain like copyright, which also requires a great degree of flexibility.

Moreover, the limited expressivity of the technical solutions currently employed makes it very difficult to accommodate copyright law into DRM systems. Consequently, DRM standards tend to follow the traditional access control approach. They concentrate their efforts in the last copyright value chain step, content consumption, and provide limited support for the other steps.

In fact, just Internet publishing risks are considered and the response is to look for more restrictive and secure mechanism to avoid access control circumvention. This makes DRM even less flexible because it ties implementations to proprietary and closed hardware and software security mechanisms.

The limited support for copyright law is also a concern for users and has been criticised, for instance by the Electronic Frontier Foundation [11]. The consequence of this lack is basically that DRM systems fail to accommodate rights reserved to the public under national copyright regimes[12].

Consequently, the DRM world remains apart from the underlying copyright legal framework. As it has been noted, this is a risk because DRM systems might then incur then into confusing legal situations. Moreover, it is also a lost opportunity because, from our point of view, ignoring copyright law is also ignoring a mechanism to achieve interoperability.

It is true that copyright law diverges depending on local regimes but, as the World Intellectual Property Organisation[2] promotes, there is a common legal base and fruitful efforts towards a greater level of copyright law worldwide harmonisation.

A new approach is necessary if we want profit from the Internet as a content sharing medium. The existence of this opportunity is clear when we observe the success of the Creative Commons initiative, whose objective is to promote content sharing and reuse thorough innovative copyright and licensing schemes.

However, despite the success of Creative Commons licenses, this initiative is not seen as an alternative to DRM. The main reason is the lack of flexibility of the available licensing terms. There are mainly six different Creative Commons licenses, all of them non-commercial, and just an informal mechanism for extension and adoption of alternative licensing schemes, CC+[3].

[2] WIPO, World Intellectual Property Organization, http://www.wipo.int
[3] http://wiki.creativecommons.org/CCPlus

Moreover, Creative Commons licenses are available in three formats: a legal version for lawyers, a more readable version for average users and as metadata for computers consumption. However, the Creative Commons metadata is not a formal representation of the licenses; it just provides a reduced set of terms for building computer-oriented licenses. There are three kinds of permissions (reproduction, distribution and derivative works), one prohibition (commercial use) and four requirements (attribution, notice, share alike and source code).

Consequently, although it is possible to provide computer support for simple services like content search, there are no mechanisms for customisation and advanced computerised support that enable an Internet-wide copyright-based alternative to DRM systems.

Formal knowledge facilitates the use of computerised means for information processing. However, at the endpoint of information processes there are human users, for which formal languages are not intended. These are users that are not experts in formal languages so they find very difficult to deal with logic or diagrammatic expressions.

Therefore, there is the need for some layer that isolates end-users from the knowledge representation language complexities. On one hand, these tend to be some kind of natural language processing or specific input forms in the knowledge acquisition end-point. On the other hand, there are specific graphical user interfaces in the knowledge presentation end-point.

In both end-points there is the need for a lot of work as these are the steps of information processing flux that require more ad-hoc development. Therefore, they constitute the bottlenecks of the whole process during the daily exploitation of knowledge-based systems.

An alternative to make knowledge arrive end-users is to use Controlled Natural Languages. Controlled Natural Languages (CNL) makes information more usable and facilitate the development of knowledge systems [13]. The need for ad-hoc knowledge acquisition and presentation means is mitigated because the acquisition and presentation functionalities are inherent to them.

However, there are still many limitations for an extensible adoption of Controlled Natural Languages. The greatest limitation we have observed is about the quality of the language sentences from the natural language point of view when the underlying knowledge structures have not been conceived for this use.

This is especially relevant in the motivating scenario of this work, the development of a Digital Rights Management (DRM) system that facilitates users deal with copyright. In this application scenario, users have extensively adopted Creative Commons licenses [14] showing that they prefer approaches based on a natural language rendering instead of those based on other approaches. However, Creative Commons lacks a computer-friendlier version that facilitates the automation of advanced DRM.

Therefore, the objective of an advanced and usable DRM system motivated the development of a pattern for natural language oriented ontology engineering. This pattern, as it is exemplified by the Copyright Ontology and the CNL-based licenses it supports, facilitates the generation of usable CNL expressions. Altogether, it facilitates the development of usable knowledge-based systems.

This works relies on existing CNL initiatives [15] as it does not propose any new alternative. On the other hand, it proposes a pattern oriented towards engineering ontologies that make CNL more natural and usable for end-users. The main contribution is instead in the ontology-engineering domain, where it proposes some guidelines geared towards producing ontologies that facilitate the generation of more usable CNL renderings.

The pattern contributed in this work concentrates on this point, i.e. to provide some patterns for ontology content creation that, additionally, facilitate the generation of usable CNL expressions. On the other hand, there is no contribution by this pattern in the ontology development process and the idea is to make these patterns compatible with the existing methodologies. Therefore, other methodologies can be reused and complemented with the proposed pattern for guided ontology content creation and usable CNL expressions generation.

2 A NL-Oriented Ontology Engineering Pattern

As it has been pointed out in the related work section, this work proposes an ontology creation pattern. This pattern must be complemented with an ontology development process and other support activities in order to guarantee that ontologies are engineered in a proper way.

Consequently, there is the need to select one of the existing ontology engineering methodologies in order to put the NL oriented pattern into practice. The Methontology [16] methodology has been chosen because it provides guidance for ontology development process but also for other support and management activities. The NL-Oriented pattern is integrated into an ontology developing process composed by the following phases: specification, conceptualisation, formalization, implementation and maintenance.

The specification phase corresponds to the pre-development aspects, where the development requirements are identified. The maintenance phase is a post-development activity, it is performed once the ontology is developed. During the conceptualisation activity, the domain knowledge is structured as meaningful models. This is the point where the proposed pattern is applied.

Moreover, if a formal language is used to build the model, it is possible to automate the formalisation and implementation activities so all the pattern work is applied just in this phase and its benefits automatically propagated though the other development activities.

Therefore, the pattern is applied during the conceptualisation activity, when the domain models are built. The pattern is applied starting with the static part of the domain at hand. The static part corresponds to the concepts called continuants or endurants [17]. Then it is time to the dynamic part, which corresponds to the concepts called ocurrents or perdurants [17]. The dynamic part is the part for which there is less proposed work in the ontology engineering area and where a greater contribution can be made in order to make the resulting knowledge structures more usable.

The objective is to facilitate the process of building a model for the dynamic ontology aspects and guarantee that this model can be translated to CNL in a more natural way, i.e. the resulting CNL expressions are more usable for users that are not formal languages experts. The proposed pattern is inspired by the way we actually model the dynamic aspects of the world using our main knowledge representation tools, i.e. natural language. Our tool for this is the verb, which models the dynamic aspects and constitutes the central point of sentences.

The objective is to apply this same pattern when modelling the dynamic aspects of an ontology. The first step is to identify the verb concepts corresponding to the ocurrents in the domain at hand, i.e. processes, situations, events, etc. These concepts will constitute the main part of the model for the dynamic part, just the same role verbs play in NL sentences.

This first step just identifies some concepts that are not enough to build complex knowledge expressions. In order to do that, the inspiration is also from how NL sentences work. In NL sentences, the verb is connected to other sentence constituents, i.e. participants, in order to build expressions that model processes, events, situations, etc. This kind of connection has been studied for long in the NL domain and a characterisation of them has been made. These connections are characterised as verb fillers called case roles or thematic roles [18].

This approach has been extensively used in the NL research domain but there is little work about applying case roles for knowledge representation. There is the FrameNet [19] initiative but it is mainly oriented towards knowledge acquisition from NL sources by semi-automatic annotation.

Two of the main proposals about the application of case roles for knowledge representation are those for Sowa [20] and Dick [21]. From these sources, a selection of case roles that can be extensively used to model the dynamic part of ontologies has been built. The contribution of this selection is that it is specially tailored to be integrated as a pattern for ontology engineering.

Table 1 shows this case roles selection, which is organised in four classes of generic case roles, which are shown at the top, and six categories, which are shown at the right. These categories correspond to verb semantic facets, not disjoint classes of verbs. Therefore, the same verb concept can present one or more of these facets. For instance, the play verb can show the action, temporal and spatial facets in a particular sentence.

Consequently, once the verb concepts have been identified, the second step of the proposed pattern corresponds to the process of determining the case roles that are necessary to build the dynamic model. Formal methods can be employed to constraint how the verb concept and the case roles are related. Therefore, this pattern allows a great range of model detail levels. Moreover, it is a very complete set of case roles. It includes all the case roles identified in the refereed bibliography and, as it is shown in the next section, it has been used during the Copyright Ontology development. During this development process no case role

	initiator	resource	goal	essence
Action	agent, effector	instrument	result, recipient	patient, theme
Process	agent, origin	matter	result, recipient	patient, theme
Transfer	agent, origin	instrument, medium	experiencer, recipient	theme
Spatial	origin	path	destination	location
Temporal	start	duration	completion	pointInTime
Ambient	reason	manner	aim, consequence	condition

Table 1. Case roles for the NL-Oriented Pattern

lack was detected and all the verb models could be built with just the case roles in Table 1.

3 The Copyright Ontology

This section details the Copyright Ontology conceptualisation activity. This activity is used as an illustrative example of the pattern presented in the previous section, which was employed in the Copyright Ontology engineering process.

The copyright domain is a complex one and conceptualising it is a very challenging task. The conceptualisation process, as it has been shown in the pattern description, is divided into two phases. The first one concentrates on the static aspects of the domain. The static aspects are divided into two different submodels due to its complexity.

First, there is the creation submodel. This model is the basis for building the conceptual models of the rest of the parts. It defines the different forms a creation can take, which are classified following the three main points of view as proposed by many upper ontologies, e.g. the Suggested Upper Merged Ontology [22]:

- **Abstract**: Work.
- **Object**: Manifestation, Fixation and Instance.
- **Process**: Performance and Communication.

A part from identifying the key concepts in the creation submodel, it also includes some relations among them and a set of constraints on how they are interrelated. More details for this point and the following steps in the conceptualisation process are available from[4].

Second, there is the rights submodel, which is also part of the static part model. The Rights Model follows the World Intellectual Property Organisation (WIPO[5]) recommendations in order to define the rights hierarchy. The most

[4] A Semantic Web approach to Digital Rights Management, http://rhizomik.net/˜roberto/thesis

[5] WIPO, http://www.wipo.int

relevant rights in the DRM context are economic rights as they are related to productive and commercial aspects of copyright. All the specific rights in copyright law are modelled as concepts. For the economic aspects of copyright there are the following rights: Reproduction, Distribution, Public Performance, Fixation, Communication and Transformation Right.

Each right governs a set of actions, i.e. things that the actors participating in the copyright life cycle can perform on the entities in the creation model. Therefore, it is time to move to the dynamic aspects of the domain. The model for the dynamic part is called the Action Model and it is built on the roots of the two previous ones.

Actions correspond to the primitive actions that can be performed on the concepts defined in the creation submodel and which are regulated by the rights in the rights submodel. For the economic rights, these are the actions:

- **Reproduction Right**: *reproduce*, commonly speaking *copy*.
- **Distribution Right**: *distribute*. More specifically *sell*, *rent* and *lend*.
- **Public Performance Right**: *perform*; it is regulated by copyright when it is a public performance and not a private one.
- **Fixation Right**: *fix*, or *record*.
- **Communication Right**: *communicate* when the subject is an object or *retransmit* when communicating a performance or previous communication, e.g. a re-broadcast. Other related actions, which depend on the intended audience, are *broadcast* or *make available*.
- **Transformation Right**: *derive*. Some specialisations are *adapt* or *translate*.

At this point we have completed the first phase of the dynamic model part, i.e. the verb concepts have been identified. They constitute the key elements in order to build knowledge expressions that represent the processes, events and situations that occur in the copyright domain.

In order to build this expression and relate the verb concepts to the other participants, i.e. concepts in the creation submodel or reused from other ontologies, it is time to complete the dynamic model and detail for each verb concept the corresponding case roles.

Due to space limitations, this section includes just the detailed model for the *Copy* action, which is formally known as *Reproduce*. However, it is commonly referred to as *Copy* and this term is the one that is going to be used in the ontology in order to improve its usability. Copies have been traditionally the basic medium for *Work* commercialisation. They are produced from a *Manifestation,* from a *Fixation* of a *Performance* or from another *Instance*. Therefore, these are the *theme* of the *Copy* verb as it is shown in Table 2.

The result is an *Instance* that is the item employed for the physical commercialisation of works, i.e. when a physical item is used as the vehicle to make the *Work* arrive to its consumers. For example, the making of copies of a protected work is the act performed by a publisher who wishes to distribute copies of a text-based work to the public, whether in the form of printed copies or digital media such as CD-ROMs.

Case role	Range	Cardinality
agent	Person (Natural or Legal)	1..N
theme	Manifestation OR Fixation OR Instance	1
result	Instance	1
pointInTime	e.g. ISO8601	1
location	e.g. ISO3166, URL, ...	1
...	...	...

Table 2. Copy case roles

The central part of figure 1 shows an example model for expression build using the proposed pattern as it is applied to the Copy verb concept. This kind of action patterns are also used to model licenses. Therefore, two additional verb concepts are identified and detailed using case roles: *Agree* and *Disagree*. They are the building block of any license. Figure 1 shows a license for the *Copy* action previously introduced. As it is shown, the *condition* case role is used in order to introduce a compensation for the agent that grants the copy action, a 3€ transfer from the granted agent.

As it can be observed in the figure, the *condition* case role is used to model the obligation deontic aspect inherent in copyright licenses. The permission and prohibition deontic aspects also present in licenses are captured by the *Agree* and *Disagree* verb concepts and their corresponding *theme* case roles.

The agreement *theme* corresponds to an implicit permission, i.e. the theme of an agreement is permitted. The *condition* relation corresponds to an obligation, i.e. in order to fulfil the theme action it is necessary to satisfy the pattern defined by the condition property object. Finally, it is also possible to model prohibitions using the *Disagree* verb concept and placing the prohibited action in the corresponding *theme*.

As a result of the Copyright Ontology development process, it has been possible to test the first objective of the proposed ontology engineering pattern. It facilitates the ontology conceptualisation because it provides a predefined pattern to face the conceptualisation process and a predefined set of constructs, the proposed case roles, which facilitate building a detailed model for the dynamic model aspects.

The other objective is to improve the usability of the expression build from the engineered ontology. This objective is based on a CNL rendering for these expressions, which is also facilitated by this pattern as it is shown in the following section.

A part from the Copyright Ontology conceptualisation presented in this section, there is an implementation[6] based on the Web Ontology Language (OWL), concretely on the Description Logic (DL) variant. This implementation can be used to develop a Semantic DRM System based on DL reasoning [23].

[6] Copyright Ontology, http://rhizomik.net/ontologies/copyrightonto

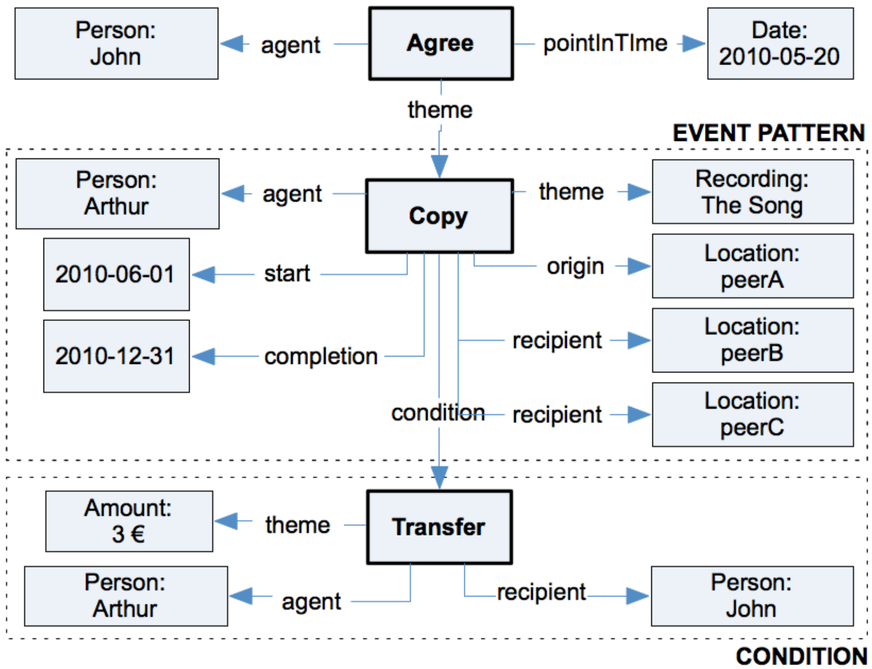

Fig. 1. Model for an agreement on a copy action pattern plus a condition

4 Mapping to Controlled NL

This section details how the pattern presented in the previous section facilitates rendering knowledge representation expressions as Controlled Natural Language. As the expressions are based on an ontology modelled from the previous NL-oriented pattern, it is easier to perform this rendering.

The mapping to CNL is simplified because it is possible to take profit from the sentence structure already present in the knowledge expression. The key point is how to render the case roles. Depending on the case role, the role filler can become the subject, the object or other component usually attached to the sentence by a preposition.

However, this is not a direct mapping between the case roles representation an NL. This would force a very detailed case-by-case procedure that would make the rendering very complicated to implement. Therefore, the approach we have undertaken is to stress the fact that we are generating CNL, which is a relaxed version of NL.

Consequently, many choices are removed and, although there is the possibility that the generated sentences are not grammatically correct. On the other hand, these limitations make the mapping much simpler.

Table 3 shows a summary of the mapping that has been currently implemented in order to perform the translation to CNL. For each case role, there are the CNL particles used to attach it to the sentence. If there is more than one choice, the table specifies the condition under with that choice is made.

First, the objective of the mapping is to get the subject from the "agent", "effector", "experiencer" or "instrument" case roles. The preference is placed in the first two and the later are chosen when none of the previous is available in the expression being rendered as CNL. If "instrument" is not performing the subject role, it is attached to the sentence using the *with* particle.

Once the subject has been placed, it is time to place the concept for the verb, just considering the concordance with the subject number, singular or plural. It is important to note that, for any case role, if there are multiple instances of the case role they are rendered together using the *and* or *or* connectives as defined in the underlying formal expression being translated to CNL.

Then it is time for the "theme" and "patient" case roles, which perform the object role in the sentence. The *that* particle is used if the theme is a subordinate sentence, i.e. the theme point to a whole sentence pattern with its own verb concept. In this case the *that* particle and the case role filler are place at the end of the current sentence, once all the other case roles have been processed. If the theme is not a verb concept, then the case role filler is directly placed after the verb.

For the rest of the case roles, the mapping choices try to consider the more common particle used to build the corresponding case role in NL. In some cases this leads to some grammatical errors for some specific verbs. Additionally, if the choice was not clear, the alternative is to use a short sentence that generalises the meaning of the corresponding case role. This is the case for the case roles in the ambient group, e.g. the "condition" case role is mapped to the connective "with condition" in the CNL sentence.

It might be the case that these choices generate sentences that are not grammatically correct and, what is more important for the work at hand, are hardly usable for end-users. In order to try to minimise this problem, a usability test has been performed. However, due to resource and time limitations, the test has concentrated on the copyright domain. Moreover, this is the scenario where this procedure is intended to be applied first.

The usability test is based on a comparative study with NL sentences capturing the intended semantics of the corresponding Copyright Ontology expressions, like the ones shown in the following examples. The comparative study is between NL and CNL sentences because the test users are not logic experts and, therefore, are not used to logic notations.

The study is based on a correlation exercise between NL sentences and the corresponding CNL ones. The latter are generated from the formal expressions corresponding to the semantics of the NL sentences. From user exercise responses it was possible to detect some problematic problems and to find more appropriate alternatives. For instance, for the case roles in the ambient group many alternatives were tested and the only solution for a direct and usable mapping

Case Role	Mapping Condition	CNL	Case Role	Mapping Condition	CNL
agent		(subject)	origin		from
effector		(subject)	start		from
experiencer	there is not agent/effector	(subject)	origin		from
instrument	there is not agent/effector/ experiencer otherwise	(subject) with	destination		to
theme	range is verb range is not verb	that (end of sentence) (object)	recipient		to
patient		(object)	result		resulting
matter		of	completion		until
medium		by	reason		with reason
pointInTime		at	manner		with manner
location		at	aim		with aim
path		through	consequence		with consequence
duration		for	condition		with condition

Table 3. Mapping case roles to CNL

was to use the connective formed by the world with and the case role name, for instance *with manner* for the "manner" case role.

In any case, in order to mitigate the user sensibility to grammatical errors and the limitations of a CNL rendering, the approach has been also to render CNL sentences in a schematic way. Instead of building a normal sentence, the CNL sentence is split up into its particles and each case role is shown in a new line. This approach has been adopted because user tests have shown that users are less sensible to grammatical errors when they face a text shown in a schematic way.

4.1 Copyright Ontology CNL Expressions Examples

The previous procedure is illustrated with the following examples. All of them are from the copyright domain because this has been the kind of expressions that have been more extensively used to test the procedure:

– In this example, the objective is to model the action corresponding to a performance of a song called "The Song" that is played by Peter. This action is first modelled using N3[7] and the concepts in the Copyright Ontology, using the "co" namespace. Then, the CNL rendering for this expression is shown using the mappings in Table 3.

```
[a co:Perform;
  co:agent :Peter;
  co:theme :The_Song].
```

"Peter perform*s* The Song".

- It is possible to model a full license. In this example, John a copy action with the condition that there is a compensatory monetary transfer to him.

```
[a co:Agree; co:agent :John; co:pointInTime ''2007-11-20'';
   co:theme

   [a co:Copy;
      co:agent :Arthur;
      co:theme :The Song;
      co:origin :peerA;
      co:destination:peerB,:peerC;
      co:start
      co:start ''2008-01-01'';
      co:completion ''2008-07-01'';
      co:condition

      [a co:Transfer; co:agent :Peter; co:recipient :John;
         co:theme [a co:Amount; rdf:value 3; cr:currency cr:Euro]
         ]]].
```

"John agree*s* **at** 2007-11-20 **that** Arthur cop*ies* The Song **from** peerA **to** peerB *or* peerC **from** 2008-01-01 **until** 2008-07-01 **with condition** Peter transfer*s* amount 3 Euro **to** John".

5 Evaluation

The Copyright Ontology and the CNL rendering are currently being evaluated in the context of the OMediaDis[8] project. The Copyright Ontology is being tested with just a limited set of licenses for a media marketplace that complement those tested in previous projects, e.g. the e-learning domain[24].

The preliminary results show that users are very pleased with the CNL rendering of the licenses in OMediaDis. In fact, the CNL rendering is not the default rendering of licenses right now. Users, when browse the license associated to a

[7] Notation3 (N3) A readable RDF syntax, http://www.w3.org/DesignIssues/Notation3.html
[8] OMediaDis, http://omediadis.udl.cat

piece of content, are shown a rendering based on HTML nested tables that show the overall structure of the license by default.

They should click an action associated to all licenses called "Read" in order to see the CNL rendering. Our experience shows that once users are told about this option, they use always as the preferred way to browse licenses because they find them more readable and easier to understand.

6 Conclusions and Future Work

There seems to be a trade-off between the level of formality of a Rights Expression Language and its usability for end-users that are not aware of the formalism employed, in our case a Web ontology. However, it is possible to reuse concepts from the natural language processing field in order to build an ontology engineering pattern that makes it possible to render this licenses using a limited natural language, called Controlled Natural Language, that makes licenses almost as appealing to end-users as the Creative Commons ones, the more widely used ones.

The pattern is inspired by how dynamic knowledge is modelled in natural language. The key issue is to use verb concepts to model processes, events, situations, etc. Additionally, case roles are also used as a knowledge representation pattern, which are used in natural language processing in order to analyse how verbs are related to the other participants in sentences.

The future work concentrates now on performing more detailed end-user test in order to improve the usability of the CNL rendering, especially when broadening the range of licences to be rendered. Currently, this approach has been tested with a set of licenses for e-learning object and in the context of the media marketplace built in the OMediaDis research project. In both scenarios, the Copyright Ontology is used as the foundation for license modelling and the same generic patterns are used to generate the CNL rendering. Consequently, the scalability of the proposal seems promising.

Another interesting aspect to explore is the performance of an ontology engineered using the NL-oriented pattern for knowledge acquisition tasks from natural language sources. At a first glance, it seems easy to reverse the CNL to case roles mapping, which might make the mapping from natural language to formal expressions based on case roles and verbs concept quite natural and direct.

Acknowledgments

The work described in this paper has been partially supported by Spanish Ministry of Science and Innovation through the Open Platform for Multichannel Content Distribution Management (OMediaDis) research project (TIN2008-06228).

References

1. Uschold, M. & Grüninger, M.: Ontologies: Principles, Methods and Applications. Knowledge Engineering Review 11(2), 93–155, 1996.,
2. García, R., A Semantic Web Approach to Digital Rights Management. Saarbrücken, Germany: VDM Verlag, 2010.
3. Lessig, L.: The Future of Ideas: The Fate of the Commons in a Connected World. Vintage, 2002.
4. Garcá, R., Gil, R.: Content Value Chains Modelling using a Copyright Ontology. Information Systems, Vol. 35, No. 4, pp. 483-495, 2010.
5. Schwitter, R.: Representing Knowledge in Controlled Natural Language: A Case Study. In M. G. Negoita, R. J. Howlett, L. C. Jain (Eds.): Knowledge-Based Intelligent Information and Engineering Systems Proceedings, Part I (pp. 711-717). LNAI 3213. Berlin, DE: Springer, 2004.
6. Rosenblatt, B.: 2008 Year in Review: Part 1. DRM Watch, December 28, 2008 http://www.drmwatch.com/drmtech/article.php/3793156
7. Wang, X., DeMartini, T., Wragg, B., Paramasivam, M., Barlas, C.: The MPEG-21 rights expression language and rights data dictionary. IEEE Transactions on Multimedia, 7(3), 408-417, 2005.
8. Iannella, R.: Open Digital Rights Language (ODRL), Version 1.1, 2002.
9. García, R., Delgado, J.: An Ontological Approach for the Management of Rights Data Dictionaries. In Moens, M., Spyns, P. (Eds.): Legal Knowledge and Information Systems. IOS Press, Frontiers in Artificial Intelligence and Applications, 134, 137-146, 2005.
10. Rosenblatt, B.: 2009 Year in Review: Part 1. DRM Watch, December 28, 2009. http://copyrightandtechnology.com/2009/12/28/2009-year-in-review-part-1/
11. Doctorow, C.: Critique of NAVSHP (FP6) DRM Requirements Report. Electronic Frontier Foundation, 2005. http://www.eff.org/IP/DRM/NAVSHP
12. Springer, M., García, R.: Promoting Music Sampling by Semantic Web-Enhanced DRM Tools. In Grimm, R.; Hass, B. (Eds.): Virtual Goods: Technology, Economy, and Legal Aspects. Nova Science Publishers, 2008.
13. Fuchs, N.E., Kaljurand, K., Kuhn, T.: Attempto Controlled English for Knowledge Representation" In Baroglio, C., Bonatti, P.A., Maluszynski, J., Marchiori, M., Polleres, A., Schaffert, S. (Eds.): Reasoning Web 2008, pp. 104–124. Springer, 2008.
14. Lessig, L-: The Creative Commons. RBL, 2003.
15. Kuhn, T., Royer, L., Fuchs, N. E., Schroeder M.: Improving Text Mining with Controlled Natural Language: A Case Study for Protein Interactions. In Proc. of the Intern. Workshop on Data Integration in the Life Sciences (DILS'06). LNBI 4075, pp. 66-81, Springer, 2006.
16. Gómez-Pérez, A., Fernández-López, M., Corcho, O.: Ontology Engineering. Advanced Information and Knowledge Processing Series. Springer, 2004.
17. Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: Sweetening Ontologies with DOLCE. In A. Gómez-Pérez, A. & R. Benjamins (Eds.): Knowledge Engineering and Knowledge Management, pp. 166-181, Springer, 2002.
18. McRae, K., Ferretti, T. R., Amyote, L.: Thematic Roles as Verb-specific Concepts. Language and Cognitive Processes Journal, 12(2), 137-176, 1997.
19. Fillmore, C. J., Johnson, C. R., Petruck, M. R. L. 2003. Background to FrameNet. International Journal of Lexicography, 16(3), 235-250.
20. Sowa, J. F. 2004. Common Logic Controlled English. Draft, 24 February 2004. Available from http://www.jfsowa.com/clce/specs.htm

21. Dick, J.P. 1991. A conceptual, case-relation representation of text for intelligent retrieval. PhD Thesis. University of Toronto, Canada.
22. Niles, I., Pease, A. 2001. Towards a Standard Upper Ontology. In C. Welty & B. Smith (Eds.) Proceedings of the 2nd International Conference on Formal Ontology in Information Systems, FOIS (pp. 2-9). Maine, USA.
23. García, R., Gil, R. 2008. Copyright Licenses Reasoning using an OWL-DL Ontology. In Breuker, J., Casanovas, P., Klein, M.; Francesconi, E. (Eds.) Law, Ontologies and the Semantic Web. Frontiers in Artificial Intelligence and Applications, Vol. 188 (pp. 145-162). Amsterdam, NL: IOS Press.
24. García, R., Pariente, T.: Interoperability of Learning Objects Copyright in the LUISA Semantic Learning Management System, Information Systems Management, Vol. 26, No. 3, 2009, pp. 252–261.