Copyright Licenses Reasoning using an OWL-DL Ontology

Roberto GARCÍA[1], Rosa GIL
Universitat de Lleida, Spain

Abstract. In order to extract the full potential from Internet-wide content sharing and reuse, the underlying copyright issues must be taken into account. The novel requirements are not satisfied by traditional Digital Rights Management. Open licensing initiatives seem more appropriate, but they lack the required computerised support. Our proposal facilitates interoperation while providing a rich framework that accommodates copyright law and copes with custom licensing schemes. It is based on the Description Logic variant of the Web Ontology Language (OWL-DL) and constitutes an ontology that conceptualises the copyright domain. The ontology provides the building blocks for flexible machine-understandable licenses and facilitates implementation because DL reasoners can be directly used for license checking. However, some preliminary transformations of the licenses models are required in order to overcome the Open World Assumption inherent in OWL-DL, which limits DL-based license reasoning.

Keywords. Digital Rights Management, Copyright, Licensing, Description Logic, Reasoning

Introduction

Traditionally, copyright management has been achieved through Digital Rights Management (DRM) systems. For instance, they have been used by record companies to protect music sold on the Internet and in enterprises in order to control content access.

DRM focuses on controlling content access, the last step in the copyright value chain, and pays little attention to the previous ones: creation, derivation, recording, communication, etc. This is enough in closed domains, like enterprise DRM or vertical content distribution channels.

However, traditional DRM is showing its limitations in Internet-wide scenarios or when it must accommodate new copyright schemes like open source or open access. For instance, a key scenario with these requirements is inter-organisational scientific and technological knowledge sharing and reuse among universities, research centres, etc.

On the other hand, there are open licensing initiatives, like Creative Commons, which show really promising results. However, they lack the required computerised support and flexibility to scale to Internet-wide copyright management.

Our proposal facilitates interoperation and automation, while providing a rich framework that accommodates copyright law and custom licensing schemes. It is based

[1]Corresponding Author: Jaume II, 69; E-25001 Lleida, Spain; E-mail: rgarcia@diei.udl.cat

on a copyright ontology, which is implemented using the Description Logic variant of the Web Ontology Language. This approach facilitates implementation because existing Semantic Web tools can be easily reused.

The rest of this paper is organised as follows. First, the next subsection explores existing initiatives and their limitations are presented, from classical and standard DRM to open access proposals like Creative Commons. Then, the Semantic Web approach to copyright-aware DRM is presented in Section 1, which is materialised in the Copyright Ontology detailed in Section 2 and implemented using Semantic Web tools as it is shown in Section 3. Finally, Section 4 presents the conclusions and the future work.

Related Work

The DRM Watch review on DRM standards [1] shows that interoperability is a key issue for DRM systems. It arises in the content distribution scenario, for instance when a user wants to consume content in any of the devices he owns, or in the organisational DRM 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. One of the main ones is ISO/IEC MPEG-21 [2], whose main interoperability facilitation component is the Rights Expression Language (REL) [3].

The REL is a XML schema that defines the grammar of a license building language, so it is based on a syntax formalisation approach. There is also the MPEG-21 Rights Data Dictionary (RDD) that captures the semantics of the terms employed in the REL, but it does so without defining formal semantics [4].

This syntax-based approach is also common to other DRM interoperability efforts and one of main causes of the lack of production implementations also observed in the DRM Watch review. Despite the great efforts in place, the complexity of the DRM 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 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 [5]. The consequence of this lack is basically that DRM systems fail to accommodate rights reserved to the public under national copyright regimes.

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 to take 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 [6], whose objective is to promote content sharing and reuse thorough innovative copyright and licensing schemes.

However, despite the success of Creative Commons licenses, who estimates more than 140 millions of works licensed under its terms, 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 no mechanism for easy extension and adoption of alternative licensing schemes.

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.

To conclude the related work overview, the closest initiatives in the Semantic Web field are the generic policy languages KAoS [7] and Rei [8]. KAoS is based on OWL and it is able to reason about policies by ontological subsumption. However, it requires some OWL-Full reasoning capabilities and its implementation is based on a theorem prover, which causes serious scalability problems. On the other hand, Rei is based on rules that overcome OWL expressivity limitations. However, this prevents it from exploiting the full potential of the OWL language. In fact, Rei rules knowledge is treated separately from OWL ontology knowledge due to its different syntactical form.

1. A Semantic Web Approach to DRM

Our proposal tries to solve the limitations observed in the current DRM and Creative Commons approaches. The underlying reason for all of them is the lack of technological tools that allow building a flexible and expressive representation framework.

[2] WIPO, World Intellectual Property Organization, http://www.wipo.int

Such framework must deal with the underlying legal framework and, simultaneously, be automated in order to benefit from computerised support. This would make possible to extract all the potential from Internet-wide knowledge sharing and reuse with the support of accurate copyright management mechanisms.

The first objective is to overcome the limitations of purely syntactic approaches, like XML, and their lack of formal semantics. The best way to formalise semantics is to use ontologies in order to build an expressive and flexible computer-supported copyright management.

Moreover, as we want a Web-wide scope, the best choice is to use an ontology language based on Web technologies. The clear choice is Semantic Web ontologies based on the OWL standard [9], which provides a set of primitives that make possible to build web-sharable conceptualisations.

The increased expressivity of web ontologies allows us to include the underlying legal framework into the formalisation and to build the rest of the system on top of it. This is a key issue because, in order to build a generic framework that facilitates interoperability, the focus must be placed on the underlying legal, commercial and technical copyright aspects.

This is the approach for the Copyright Ontology[3], detailed in the following section. The expressiveness and generality of the resulting conceptualisations allows coping with the shortcomings of existing approaches and, additionally, the ontology can be used as an interoperability facilitator for the main DRM standards [10].

The ontology is implemented as an OWL Web ontology based on the Description Logic (DL) variant, OWL-DL. This implementation facilitates DRM systems development as license checking is implemented using existing Semantic Web reasoners.

To the best of our knowledge, there is just one other ontological framework for DRM, OntologyX[4]. However, it is a commercial product for which there is little publicly available information. In any case, from the available information, it is clear that OntologyX concentrates on the kind of actions that can be performed on governed content and it does not take into account the underlying legal framework. Moreover, it currently lacks formal semantics and can be seen more like a rights dictionary than as a fully-fledged ontology.

2. The Copyright Ontology

The copyright domain is quite complex so we face its conceptualisation in three phases. Each phase concentrates on a part of the whole domain. First, the objective is the more primitive part, the Creation Model. Second, there is the model for the rights part, the Rights Model, and finally a model for the available actions, the Action Model, which is built on top of the two previous ones.

2.1. Creation Model

The Creation Model conceptualises the different forms a creation can take, which are classified depending on the classical top ontological points of view [11,12,13]:

[3] Copyright Ontology, http://rhizomik.net/ontologies/copyrightonto
[4] OntologyX, http://www.ontologyx.com

– **Abstract**: something that cannot exist at a particular place and time without some physical encoding or embodiment.
  - **Work**: is a distinct intellectual or artistic creation. It includes literary and artistic works, music, pictures and motion pictures, but also computer programs or compilations, like databases.
– **Object**: corresponds to the class of ordinary objects and includes digital objects.
  - **Manifestation**: the materialisation of a work in a concrete medium, a tangible or digital object.
  - **Fixation**: the materialisation of a performance in a concrete medium, a tangible or digital object.
  - **Instance**: the reproduction, copy, of a manifestation, a fixation or another instance.
– **Process**: something that happens and has temporal parts or stages.
  - **Performance**: the expression in time of a work. Performers or technical methods might be involved in the process.
  - **Communication**: the transmission of a work among places at a given time. It is a process performed when the public is not present at the place and or time where the communication originates. It includes broadcasts, i.e. one to many, but also communications from a place and at a time individually chosen.

There are many relations among the different forms a creation can take during its life cycle, see Figure 1, as it evolves from an abstract idea, i.e. a *Work*, towards something that can be consumed by end users, e.g. and *Instance* or a *Communication* or *Performance*.

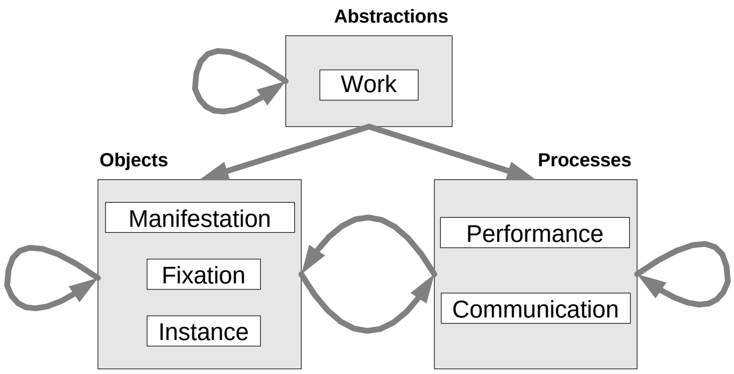

Figure 1. Creation model showing the different views on creation

This relations are named following the same pattern, e.g. for a *Manifestation* there is the relation *isManifestationOf*, which relates it to the original *Work* that it materialises, and the reverse relation *hasManifestation*, that relates a *Work* to all its manifestations.

2.2. Rights Model

The Rights Model follows the World Intellectual Property Organisation recommendations. It includes economic plus moral rights, as promoted by WIPO, and copyright related rights, see Figure 2.

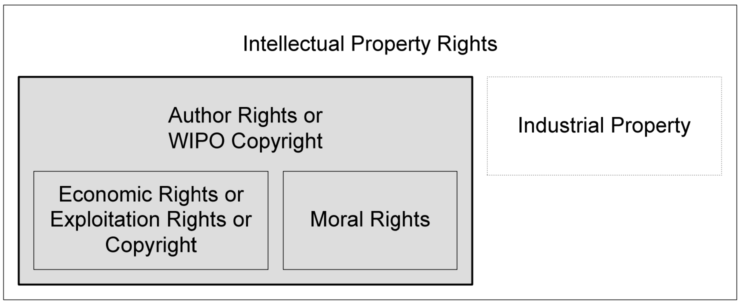

Figure 2. The Rights Model in the Copyright Ontology

The most relevant rights in the context of DRM systems are economic rights as they are related to the production and commercial aspects of copyright. Reproduction, Distribution, Public Performance, Fixation, Communication and Transformation Right are the economic rights. The Rights Model in the Copyright Ontology provides the following hierarchy of copyright rights:

- **Economic Rights**:
  - Reproduction Right
  - Distribution Right
  - Public Performance Right
  - Fixation Rights (Sound Record and Motion Picture Rights)
  - Communication Rights (Broadcasting and Making Available Rights)
  - Transformation Rights (Adaptation and Translation Rights)

There are also the moral rights, which are always held by the creator and cannot be commercially exploited, and the related or neighbouring rights, the rights of the other actors also involved in the exploitation of works: performers, producers and broadcasters.

- **Moral Rights**:
  - Attribution Right
  - Integrity Right
  - Disclosure Right
  - Withdrawal Right
- **Related Rights**:
  - Performers Rights
  - Phonograms Producers Rights
  - Broadcasters Rights

2.3. Action Model

The last model, the Action Model, corresponds to the primitive actions that can be performed on the concepts defined in the Creation Model, as it is shown in Figure 3. Actions are regulated by the rights in the Rights Model. For the economic rights, these are the governed actions:

- **Reproduction Right**: *reproduce*, commonly speaking *copy*.
- **Distribution Right**: *distribute*. More specifically *sell*, *rent* and *lend*.
- **Public Performance Right**: *perform*; it is regulated 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*.

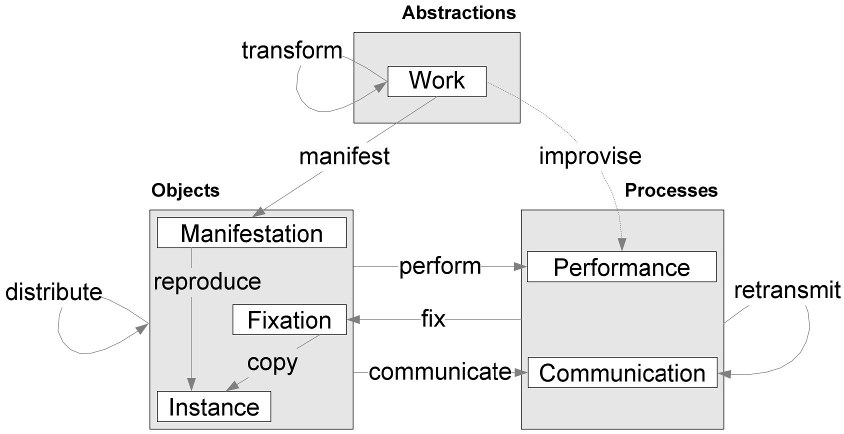

Figure 3. Relations between the Action and Creation Models

One of the biggest criticisms against DRM is that they do not respect some special permissions that many copyright legal systems provide to end-users. These permissions are commonly called fair use, fair dealing or user rights. Although some of them are referred to as rights, e.g. the right to quote, they constitute exceptions to copyright and should be considered as end-user privileges and not rights.

These privileged actions, normally restricted by copyright, may be done without the authorization of the copyright owner in circumstances specified in the law. Moreover, these exceptions do not mean that the exceptional use is always free. Some require the user to pay a compensation. For instance, in some countries, there are levies on digital recording equipment and media.

These are the main copyright exceptions:

- **Quotation Right**: quote, a limited extent *reproduce* action of a source protected work, which is clearly mentioned.
- **Education Right**: educational act, any *reproduce*, *communicate* or *perform* action with educational or research purposes.
- **Information Right**: *inform*, any copyright governed act with informative purposes.
- **Official Act Right**: official act, any copyright governed act that is part of an official act.
- **Private Copy Right**: reproduce privately, a *reproduce* act that produces a reproduction solely for private consumption.
- **Parody Right**: parody, any copyright governed act with parody or caricature purposes.
- **Temporary Reproduction Right**: reproduce temporally, a *reproduce* act that produces a temporal reproduction.

The action concepts are complemented with a set of relations that link them to the action participants. This set is adopted from the linguistics field and it is based on case roles [14]. The case roles in the Action Model are shown in Table 1.

Table 1. Action Model case roles

	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

The previously introduced pool of primitive actions and case roles allows building models for events and value chains in the copyright domain. For instance, Figure 4 shows how we can build a model for the value chain of serials adapted from literary works.

First, the creator adapts the original literary work, e.g. Alexandre Dumas' "The Count of Monte Cristo", in order to produce a serial. The resulting adaptation is realised as a script that is performed by some actors, e.g. Gerard Depardieu, and recorded into a motion picture. This motion picture is finally broadcasted to users who can tune the resulting communication.

This is just the skeleton of the value chain. In order to give a more detailed model, each step in the value chain should be modelled as an event for the corresponding action and associated participants through case roles.

However, the objective is not just to model the actual events that capture the life cycle of a given creation. Prior to these events, licenses among the involved parties are established in order to govern the value flux. Consequently, the ontology must be enriched with permissions, prohibitions and obligations [15].

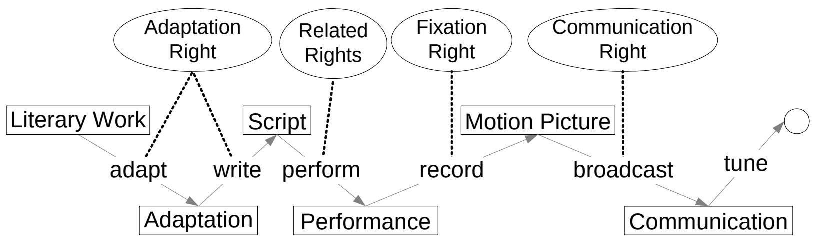

Figure 4. Literary works adapted to serials value chain

2.4. Modelling Copyright Licenses

Copyright provides a legal framework that governs creations life cycles and tries to assure a fair compensation for all the involved parties, from authors to consumers. Copyright licenses are built on top of this legal framework and establish the terms for concrete interaction among these parties.

Licenses should capture the obligations, permissions and prohibitions that make sense in the copyright domain. The semantics of the license terms are captured by the ontology described so far, but it lacks the terms that capture the semantics of obligations, permissions and prohibitions.

In order to produce a homogeneous and usable conceptualisation, we have incorporated this terms in the ontology using the concepts that capture the semantics of obligations, permissions and prohibitions as they appear in contracts from a natural language point of view, i.e. using the corresponding actions and case roles (e.g. the verb to agree on a specific action as the natural way to model a permission).

The additions are detailed next and they are related to a generic contract modelling language, the Business Contract Language (BCL) [16], in order to illustrate how these additions make the Copyright Ontology a copyright contracts modelling tool. In the following subsections, each BCL building block is introduced and then related to its Copyright Ontology counterpart.

2.4.1. Roles

The simplest building blocks in BCL are roles, e.g. *Purchaser*, which are captured in a generic way by the Copyright Ontology case roles. For instance, there is not a specific *Purchaser* case role but it is implicit in the *agent* case role when applied to a *Purchase* action.

2.4.2. Event Patterns

BCL uses event patterns as the way to state what is obliged, permitted or prohibited by a contract; they are referenced from policies that establish their modality. They are also naturally captured by the ontology terms described so far. The proposed actions and case roles are used to model event patterns in the copyright domain.

For instance, Figure 5 shows a pattern for all copy events in a Peer to Peer network performed by agent "granted" who copies "content01" from "PeerA" to two peers from the set "PeerB, PeerC, PeerD" at any time point six months after "2006-01-01".

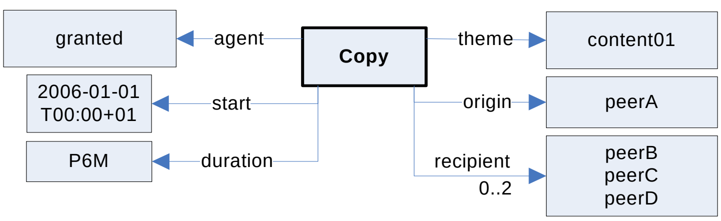

Figure 5. Pattern for a copy action in a P2P scenario

2.4.3. Modalities

Then, there are the terms to state the modality of these event patterns in copyright contracts. BCL defines explicitly the modalities using the Obligation, Permission and Prohibition terms. The Copyright Ontology does the same but in an implicit way, following the same "action plus case roles" approach used for event patterns. Additional classes and relations are added in order to attach modalities to event patterns. The objective is to state that the set of actions corresponding to the pattern is permitted, obliged or prohibited, depending to the particular construct that is attached to the pattern, as it is detailed in the next subsections.

2.4.3.1. Permissions

BCL Permissions are captured by a new action class, *Agree*, and the permitted pattern is linked using the *theme* case role, whose semantics are to point to the object of an action.

Following with the example in Figure 5, in order to authorise the pattern that it models, an agreement like the one shown in Figure 6 can be modelled. The agreement between "granter" and "granted" in the upper part of authorises the pattern pointed by the *theme* case role, the previous P2P copy pattern at the centre of the figure.

2.4.3.2. Obligations

BCL Obligations are captured in the copyright contracts as event patterns that must be satisfied at some time point after the event pattern that triggers the obligation is exercised. They are modelled using the *consequence* case role that links the triggering pattern to the one that is obliged.

For instance, in the bottom part of Figure 6 it is stated that, if the copy action is exercised, the consequence is that the "granted" agent must transfer 3 Euros to the "granter" agent before 24 hours from the copy action.

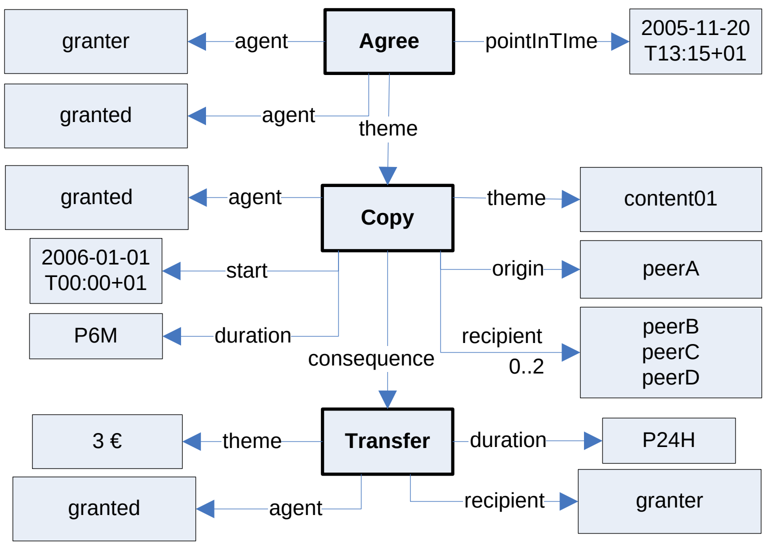

Figure 6. Agreement that permits the P2P copy pattern whose consequence is an economic obligation

2.4.3.3. Prohibitions

BCL Prohibitions are captured by another action, *Disagree*. Like for the *Agree* action, the *theme* case role is used to link it to the object of the action, in this case to the pattern that is prohibited.

For instance, in the previous scenario, the contract might also state that it is forbidden that the "granted" agent changes "content01" using a *Disagree* pattern with the corresponding *Transform* action pattern as its *theme*.

2.4.4. Guards

BCL Guards are patterns that must be satisfied in order to activate the evaluation of another event pattern, thus acting as a precondition. The *condition* case role is used to model guards. It is applied to the pattern that is guarded and it links to the pattern that establishes the precondition. The approach is similar to the obligation case captured by the *consequence* case role but, in this case, the *condition* case role establishes an a priori condition.

For instance, in the P2P scenario the *Copy* pattern might by guarded by a *Transfer* one that requires that the "granted" agent makes a 1 Euro prepayment to the "granter" agent before the former can excise the permitted P2P Copy action.

3. OWL Implementation

The previous conceptualisation is just an abstraction of the copyright domain. An implementation is required if we want to use it to build a computerised copyright management system. The Semantic Web approach is also productive in this respect because existing tools can be used to make the implementation quite straightforward.

The ontology has been implemented using the DL variant of the Web Ontology Language (OWL-DL), which is constrained in order to be managed by Description Logic (DL) reasoners. Such reasoners guarantee that OWL-DL ontologies can be put into practice, i.e. reasoned over, in a decidable and tractable way.

Existing DL reasoners are used to automatically check if actions on copyrighted content are authorised or not. As it has been shown, licenses are composed of *Agree* or *Disagree* actions, linked through a *theme* relation to patterns of actions that are correspondingly authorised or forbidden.

The pattern is implemented as an OWL class made up from the combination of classes for actions, e.g. *Copy* or *Access*, and a set of OWL Restrictions. Each restriction defines a constraint on how members of the class, the domain, are related through the specified property to other ones, the range class. The available restrictions are:

- **allValuesFrom**: all the values for the range of the restricted property must pertain to the given class. For instance, all values of the *agent* relation must pertain to the *Publisher Subscribers* class or, for the *pointInTime* relation, to the time range [2007/01/01–2007/06/30]. In order to support the later, custom datatypes reasoning is required [17].
- **someValuesFrom**: there is at least one value that pertains to the given range class.
- **hasValue**: the range is limited to a specific individual, not a class of them. For instance, the *theme* of a *Copy* action must be the individual "doi:10.1032/…".
- **cardinality**: this restriction limits the number of individuals that can be connected through the restricted property. A maximum, minimum or exact

cardinality can be defined. For instance, the *recipients* of an action can be limited to just two individuals.

Restrictions are combined using the intersection, union and complement logical operators in order to compose the patterns of actions. For instance, Figure 7 shows the conceptual model for a license that combines commercial and open access terms.

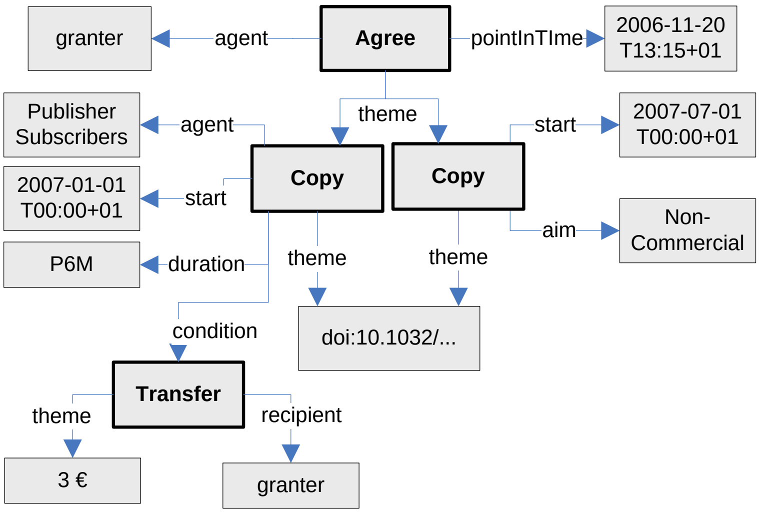

Figure 7. Agreement on a copy action under commercial and open access terms

The upper part shows an *Agree* that permits two *Copy* patterns, connected through the *theme* relation. The one of the left grants "Publisher Subscribers" to copy some content identified by a DOI at any time point six months after 2007-01-01. Any attempt to exercise this action pattern is subject to a commercial condition, a compensation of 3€. On the other hand, the *Copy* pattern on the right grants anyone to copy the same content, once the period of six months is surpassed, if the aim is non-commercial.

The constraints on the kinds of actions, their agents, time points, etc. are then implemented using OWL Restrictions, which are combined using the logical operators in the OWL language. Figure 8 shows the pattern build up from the combination of such kind of restrictions for the example presented in Figure 7. For the set of all copy actions on "doi:10.1032/…", the light grey area, two subsets are selected and their union constitutes the licensed actions pattern, the dark grey areas.

As it can be seen in Figure 8, each intersected restriction reduces the set of actions. For instance, the non-commercial pattern does not include any restriction on the agent of the action. Consequently, the licensed actions set includes any non-commercial copy action performed by anyone later than 2007-07-01. Table 2 shows the DL notation for the class definition that models the commercial copy pattern.

Table 2. OWL-DL Class for the commercial copy action pattern

Pattern $\equiv$	Copy $\sqcap$	(1)
	$\forall$pointInTime.$\geq$ 2007-01-01T00:00:00, $\leq$ 2007-06-30T23:59:59 $\sqcap$	(2)
	$\forall$agent.PublisherSubscribers $\sqcap$ ($\geq$ 1 agent) $\sqcap$	(3)
	$\exists$theme.{urn:doi.10.1032/…} $\sqcap$ ($\leq$ 1 theme)	(4)

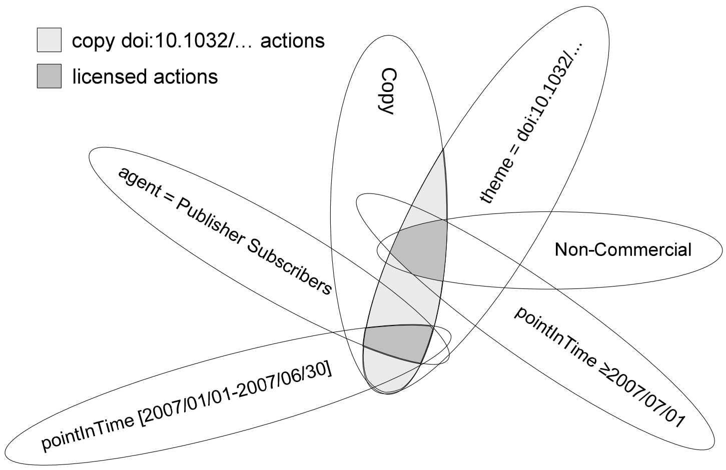

Figure 8. Building an action pattern as an intersection of restrictions

Each intersected restriction reduces the initial set of actions, which corresponds to all the *Copy* actions. First, (2) models the time range as a restriction on the *pointingTime* case role to a custom datatype. This case role is a functional property so no additional constraints on cardinality are required. The last constraints, (3) and (4), restrict the range of *agent* to one or more instances of the "PublisherSubscribers" class and *theme* to just the instance "urn:doi.10.1032/…".

From this point, the implementation is quite straightforward. DL reasoners are specially suited for classifying individuals into classes when the latter are based on necessary and sufficient conditions. They can answer if an individual, considering its relations to other individuals and attribute values, satisfies all the restrictions of a class pattern and, thus, can be classified as an instance of that class.

In the context of the Copyright Ontology implementation, this functionality is used to check if a particular action, modelled as an individual, is allowed or not by a license. This corresponds to the fact that the action individual is classified into a class pattern that is the *theme* of an *Agree*. Another reading is that the license agrees on performing a set of actions that includes the requested one.

However, before the action is authorised, it is also necessary to check that any existing *condition* is met and that there is not any disagreement on the action. The DL reasoner is also useful for this part. It is checked if the precondition pattern is instantiated, so the precondition is satisfied, and that the checked action is not classified into a class pattern that is the theme of a *Disagree*.

To sum up, it is checked that there is an agreement on the action and no disagreement, and that the precondition is satisfied. This behaviour allows modelling complex licenses, revocation and avoiding the Open World Assumption inherent to OWL-DL as it is detailed in the next section.

3.1. Overcoming the Open World Assumption

The main problem of the OWL-DL implementation presented in the previous section is the Open World Assumption (OWA). This problem arises when the DL reasoners are trying to classify a given instance into the classes for license patterns. It can be said that the reasoner is very "conservative" as, although the necessary and sufficient conditions are met, it will not classify an instance into a class if new facts can make it retract from this decision.

In some cases this is the desired behaviour but this is not the case for this license checking implementation. The intention is to make a local Close World Assumption and make a decision on the currently available facts as the outcome is to decide if the current action should be authorized or not at check time.

There are some OWL-DL restriction primitives that lead to OWA problems:

- **maxCardinality** ($\leq$ n): the reasoner is conservative with this restriction as, although the cardinality restriction might be satisfied at a given time point, new facts can make the cardinality greater than n, i.e. ($>$ n). The cardinality restriction is also affected as it is the conjunction of *maxCardinality* and *minCardinality* restrictions.
- **allValuesFrom** ($\forall$R.C): the situation in this case is that, although at the current time all the values for the R property are in the C class, in the future, there might be new facts that involve R with a value not in C, i.e. R.($\neg$C).

On the other hand, other OWL-DL restrictions, or their combination, are not affected by the OWA and thus do not affect the license checking implementation. Some of them are:

- **minCardinality** ($\geq$ n): there is no OWA problem here as once the reasoner can check that the cardinality is equal or greater than n, i.e. ($\geq$ n), new facts cannot make this inference false, i.e. ($<$ n).
- **someValuesFrom** ($\exists$R.C): once there is some R whose value is in C, new facts cannot make that there does not exist some R with a C value, i.e. $\neg$($\exists$R.C). Therefore, there is not an OWA problem here.
- **FunctionalProperty** R: this constraint makes an *allValuesFrom* restriction OWA insensitive.

There are many ways to overcome OWL-DL's OWA through epistemic operators [18] and non-monotonic OWL extensions [19]. However, in these cases it is necessary to get outside standard OWL-DL and, what is even more inconvenient, there are just preliminary implementations of these approaches.

In order to implement an OWL-DL based license checker, we have adopted a more pragmatic approach, which does not require additional language constructs neither a reasoner different from the existing OWL-DL ones. Figure 9 illustrates this approach. As it can be observed, a *maxCardinality* restriction defines a set of accepted cardinality values, e.g. from zero to two. However, as new facts are known under an OWA, instances previously classified into this restriction can "get out" of the corresponding set.

On the other hand, it can be observed that a *maxCardinality* restriction ($\leq$ n) has an opposite set corresponding to the *minCardinality* restriction ($\geq$ n+1). As it has been shown, the *minCardinality* construct is not affected by OWA. Therefore, the idea is to take profit from this fact and make the reasoner look for the opposite set, the OWA-insensitive one, and check that it is not satisfied. This implies that the reverse is satisfied by the current set of facts at hand and overcomes the OWA assumption that makes the reasoner not able to infer that. The same applies to the *allValuesFrom* restrictions and the reversed *someValuesFrom* restrictions.

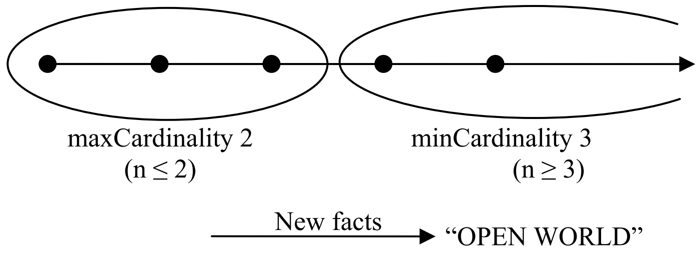

Figure 9. Opposed OWA-sensitive maxCardinality to OWA-insensitive minCardinality

Therefore, the approach is to negate the restriction and to undo this at the metalevel, i.e. to check outside the DL reasoner that the negated restriction is not satisfied and thus it can be inferred that the original one does. The negation is modelled at the metalevel using the *Disagree* class, which is the opposite of the *Agree* class.

Therefore, the *allValuesFrom* restriction on the *theme* class of an *Agree*, i.e. $\forall R.C$, is converted into the reversed *someValueFrom* restriction on the *theme* class of a *Disagree*, i.e. $\exists R.\neg C$. On the other hand, the *maxCardinality* restriction on the *theme* class of an *Agree*, i.e. $(\leq n\ R)$, is converted into a *minCardinality* restriction on the *theme* class of a *Disagree*, i.e. $(\geq n{+}1\ R)$.

The previous method is applied to class patterns and all the OWA-sensitive constructs are moved to a new class pattern which is disagreed in order to model the metalevel negation. This new class results from the disjunction of all the transformed restrictions and is intersected with the original pattern, which is now composed by just the OWA-insensitive restrictions and is what remains as the subject of the *Agree*.

Continuing with the commercial copy pattern described in Table 2, the following example illustrates this mechanism. Table 3 shows the class patterns that overcome the OWA and result from the previous transformation. Note that *pointInTime* is defined as a functional property in the Copyright Ontology so it is not affected by the OWA and remains unchanged (2), like the other OWA-insensitive constructs (3). All of them build *Pattern'* that corresponds to the OWA insensitive part of the original *Pattern*, i.e. the pattern that is agreed.

Pattern' ≡	Copy ⊓	(1)
	$\forall$pointInTime.≥ 2007-01-01T00:00:00, ≤ 2007-06-30T23:59:59 ⊓	(2)
	(≥ 1 agent) ⊓ $\exists$theme.{urn:doi.10.1032/…}	(3)
Pattern'' ≡	Pattern' ⊓	(4)
	($\exists$agent.¬PublisherSubscribers ⊔	(5)
	(≥ 1 theme))	(6)

Table 3. OWA-insensitive classes for the commercial copy action pattern in Table 2

On the other hand, *Pattern''* contains the transformed OWA sensitive constructs in the original *Pattern*, i.e. the disagreed pattern. There is the *someValuesFrom* restriction (5) corresponding to the *allValuesFrom* restriction (3) in Table 2 and the *minCardinality* restriction (6) corresponding to the *maxCardinality* one (4) in Table 2.

The combination of both patterns, the first the *theme* of an *Agree* and the second of a *Disagree*, ends up building a pattern like the one shown in Figure 10. The set of actions that is authorised corresponds to the darker part, i.e. *Pattern'* minus *Pattern''*.

The figure also shows three example instances, in N3 notation, situated into the class patterns under which they are classified. Consequently, just instances of *Pattern'* that are not instance of *Pattern''* are authorised. The values that make instances to not be classified into *Pattern'* are highlighted using bold letters. It is assumed that “Roger” is an instance of *PublisherSubscribers* and that “Matt” does not.

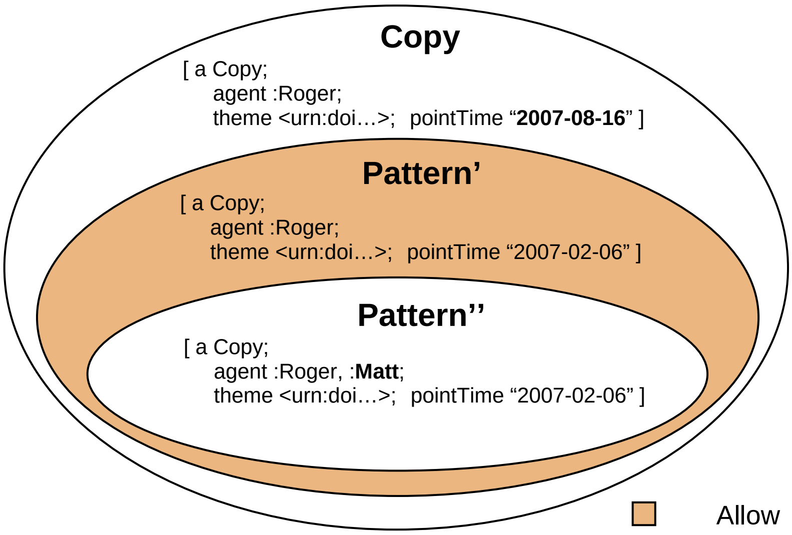

Figure 10. Interpretation of the class patterns in Table 3

3.2. Creative Commons Scenario

As it has been pointed out in the introduction, one of the more successful copyright licensing approaches is Creative Commons (CC). However, the machine-readable version of CC licenses models a very limited part of the licenses semantics, i.e. many details in the human-readable version are not captured in the computer one.

Consequently, current CC licenses are just appropriate for license search computerised support. In order to enable more sophisticated services for CC licenses, we have modelled them using the Copyright Ontology. Thus, we can make their semantics explicit and propagate the previously described license reasoning capabilities to the enormous pool of CC licenses.

As the semantics are implicit in the human-readable version, it has been necessary to interpret and manually map the licenses to Copyright Ontology concepts. Fortunately, there is a limited set of predefined licenses[5] so this is an easily affordable process.

For instance, Table 4 presents the model for the “Attribution Share-Alike” (by-sa) CC license. This license allows licensees to remix, tweak, and build upon the copyrighted work even for commercial reasons, as long as they credit the author and license their new creations under identical terms.

Consequently, in order to license under these terms any manifestation of “myWork”, the three “By-Sa” class patterns are granted. The first one, “By-Sa_1”, lets licensees to copy, distribute, communicate and make available any manifestation of “myWork” if they give credit to the author. The second one, “By-Sa_2” authorises

[5] Creative Commons Licenses 3.0, http://creativecommons.org/licenses

derivations of the work and the third one, "By-Sa_3" establishes the same terms than for the original work for those derivations.

$$\textbf{By-Sa\_1} \equiv (\text{Copy} \sqcup \text{Distribute} \sqcup \text{Communicate} \sqcup \text{MakeAvailable}) \sqcap \forall \text{theme.MyManifestation} \sqcap (\geq 1\ \text{theme}) \sqcap \exists \text{condition.AttributeMe}$$

$$\text{MyManifestation} \equiv \text{Manifestation} \sqcap \forall \text{isManifestationOf.}\{\text{myWork}\}$$
$$\text{AttributeMe} \equiv \text{Attribute} \sqcap \exists \text{theme.}\{\text{myWork}\} \sqcap \exists \text{recipient.}\{\text{me}\}$$

$$\textbf{By-Sa\_2} \equiv \text{Derive} \sqcap \exists \text{theme.}\{\text{myWork}\} \sqcap \forall \text{result.NewManifestation} \sqcap (\geq 1\ \text{result})$$

$$\text{NewManifestation} \equiv \text{Manifestation} \sqcap \forall \text{isManifestationOf.NewWork} \sqcap (\geq 1\ \text{isManifestationOf})$$
$$\text{NewWork} \equiv \text{Work} \sqcap \exists \text{isDerivationOf.}\{\text{myWork}\}$$

$$\textbf{By-Sa\_3} \equiv (\text{Copy} \sqcup \text{Distribute} \sqcup \text{Communicate} \sqcup \text{MakeAvailable}) \sqcap \forall \text{theme.NewManifestation} \sqcap (\geq 1\ \text{theme}) \sqcap \exists \text{condition.AttributeMe}$$

Table 4. Copyright Ontology model for the CC by-sa license

4. Conclusions and Future Work

We are not profiting from the full potential of Internet-wide content sharing and reuse because the underlying copyright issues are not made explicit and dealt with. Instead, the reaction is to protect content using security mechanisms that limit the possibilities.

A good example of the potential of a less restrictive approach is Creative Commons licensing schemes for open access and reuse of content. However, Creative Commons does not constitute and alternative to DRM. It lacks the required flexibility to incorporate additional license terms, like commercial ones, and advanced computerised support.

Our semantic web approach to copyright management constitutes an alternative. It provides an expressive conceptual framework, the Copyright Ontology, which provides the building blocks for flexible machine-understandable licenses.

Altogether, it constitutes a tool that helps people state the copyright conditions for the content they share and how it might be reused. A way to build an Internet-wide licensing network adapted to particular needs: commercial or non-commercial, open or closed access, reusable share-alike content, etc.

Moreover, thanks to its OWL-DL implementation and mechanisms to overcome the Open World Assumption, it can be put into practice quite easily by using existing DL reasoners. License reasoning allows checking if a particular action is granted by a pool of licenses. This capability can be propagated to existing license modelling languages, like DRM standards [10] or Creative Commons licenses.

From these results, future work focuses on combining the DL layer, which deals with patterns and events, with a rule-based metalevel. Currently, the metalevel is implemented procedurally, which is possible due to the limited range of interpretations of the DL classifications. However a metalevel implementation based on Semantic

Web rules would make available a greater level of flexibility and new functionalities. For instance, rules can facilitate incorporating penalties into the system, i.e. obligations that take place when obligations are violated [20]. Currently, obligations are just monitored in order to detect violations.

References

[1] Rosenblatt, B.: “2005 Year in Review: DRM Standards”; DRM Watch, January 2, (2006) http://www.drmwatch.com/standards/article.php/3574511

[2] de Walle, R.V., Burnett, I.: “The MPEG-21 Book”; John Wiley & Sons, UK (2005)

[3] 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 (2005), 408-417

[4] García, R., Delgado, J.: “An Ontological Approach for the Management of Rights Data Dictionaries”; In Moens, M., Spyns, P. (ed.): “Legal Knowledge and Information Systems”; IOS Press, Frontiers in Artificial Intelligence and Applications, 134 (2005) 137-146

[5] Doctorow, C.: “Critique of NAVSHP (FP6) DRM Requirements Report”; Electronic Frontier Foundation (2005) http://www.eff.org/IP/DRM/NAVSHP

[6] Lessig, L.: “The Future of Ideas: The Fate of the Commons in a Connected World”; Vintage (2002)

[7] Uszok, A., et al.: “KAoS policy management for semantic web services”; IEEE Intelligent Systems, 19, 4 (2004), 32-41

[8] Kagal, L.: “A Policy Based Approach to Governing Autonomous Behavior in Distributed Environments”; PhD Thesis, University of Maryland, Baltimore County, USA, (2004)

[9] McGuinness, D.L., van Harmelen, F.: “OWL Web Ontology Language Overview”; W3C Recommendation 10 February 2004 (2004) http://www.w3.org/TR/owl-features

[10] García, R., Gil, R., Delgado, J.: “A web ontologies framework for digital rights management”; Journal of Artificial Intelligence and Law, 15, 2 (2007), Online First

[11] Niles, I., Pease, A.: “Towards a Standard Upper Ontology”; In Welty, C.; Smith, B. (eds.): “Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS)”; ACM Press, New York (2001) 2-9

[12] Hoekstra, R., Breuker, J., Bello, M. D., Boer, A.: “The LKIF Core Ontology of Basic Legal Concepts”; In P. Casanovas, M. A. Biasiotti, E. Francesconi, M. T. Sagri,(eds.) Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques, LOAIT’07 (2007)

[13] Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: “Sweetening Ontologies with DOLCE”; In Gómez-Pérez, A., Benjamins, R. (eds.): “Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web”; 13th International Conference, EKAW’02 (2002), 166-181

[14] Sowa, J.F.: “Knowledge Representation. Logical, philosophical and computational foundations”; Brooks Cole Publishing Co. (2000)

[15] Lee, R.M.: “A Logic Model for Electronic Contracting”; Decision support systems, 4 (1988), 27-44

[16] Linington, P. F., Milosevic, Z. , Cole, J., Gibson, S., Kulkarni S., Neal S.: “A unified behavioural model and a contract language for extended enterprise”; Data & Knowledge Engineering, 51, 1 (2004), 5-29

[17] Pan, J.Z., Horrocks, I.: “OWL-Eu: Adding customised datatypes into OWL”; Web Semantics, 4, 1 (2006), 29-39

[18] Grimm, S., Motik, B.: “Closed World Reasoning in the Semantic Web through Epistemic Operators”; Proc. OWL Experiences and Direction Workshop (2005)

[19] Katz, Y., Parsia, B.: “Towards a Nonmonotonic Extension to OWL”; Proc. OWL Experiences and Direction Workshop (2005)

[20] Grosof, B.N., Poon, T.C.: “SweetDeal: Representing Agent Contracts with Exceptions Using Semantic Web Rules, Ontologies and Process Descriptions”; International Journal of Electronic Commerce, 8, 4 (2004), 61-97