

Content Value Chains Modelling using a Copyright Ontology

Roberto García¹, Rosa Gil
Universitat de Lleida
Jaume II, 69, Lleida, Spain
{rgarcia,rgil@diei.udl.cat}

Abstract. Existing Digital Rights Management systems, initiatives like Creative Commons or research works as some digital rights ontologies provide limited support for content value chains modelling and management. This is becoming a critical issue as content markets start to profit from the possibilities of digital networks and the World Wide Web. The objective is to support the whole copyrighted content value chain across enterprise or business niches boundaries. Our proposal provides a framework that accommodates copyright law and a rich creation model in order to cope with all the creation life cycle stages. The dynamic aspects of value chains are modelled using a hybrid approach that combines ontology-based and rule-based mechanisms. The ontology implementation is based on the Web Ontology Language (OWL-DL) and Description Logic reasoners are directly used for license checking. On the other hand, for more complex aspects of the dynamics of content value chains, rule languages are the choice.

Keywords. Digital Rights Management; Copyright; Content Value Chain; Ontology; Reasoning.

1. Introduction

Nowadays, content value chains are managed by proprietary Digital Rights Management (DRM) systems. Some standardisation attempts have been carried out, with some successful experiences in the mobile phones domain. However, these standards provide little support for the whole value chain as they focus on content distribution to mobile phones.

This is due to the fact that 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 might be enough in closed domains and vertical markets, like enterprise DRM or specific content distribution channels.

Internet-wide content value chains, where all the participants can interact with content in a distributed and open manner, are not supported by such systems. DRM does not incorporate the copyright notions that govern these value chains and lacks the mechanisms that make possible to accommodate new copyright schemes like open source or open access.

On the other hand, there are open licensing initiatives, like Creative Commons [1], which have been very successful for non-commercial licensing schemes. However,

¹ Corresponding author: telephone +34-973-702-742, fax +34-973-702-702

they lack the required computerised support and flexibility to scale to Internet-wide copyright management.

Our contribution tries to leverage DRM systems, also called Digital Restriction Management systems [2], to copyright management systems, which support the whole value chain, from creators to consumers, and build on top of copyright law. The proposal is based on a copyright ontology that provides the building blocks and restrictions that make it possible to model content value chains with a level of detail and formalisation that facilitates developing copyright management systems.

This approach is related with other ontological approaches to DRM. However, it contributes the copyright dimension, which is commonly ignored, and support for the whole value chain, which is difficult to attain if the underlying legal framework is not taken into account. Moreover, the copyright ontology gains relevance due to ISO/IEC MPEG recent interest on ontologies [3] as a mechanism to facilitate the implementation of existing standards like MPEG-21 [4] and provide additional content services.

The rest of this paper is organised as follows. First, Section 2 explores related work and current initiatives. Then, the copyright ontology is presented in Section 3 and its dynamic aspects in Section 4. Finally, Section 5 presents the conclusions and the future work, together with an analysis of how the Copyright Ontology fits MPEG Media Value Chain Ontology requirements.

2. Related Work

The DRM Watch review on DRM standards [5] 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, whose main interoperability facilitation component is the Rights Expression Language (REL) [6]. The REL is a XML Schema that defines the grammar of a license building language. 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 [7].

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 existing ones, for instance the OMA² version of ODRL [8], are implementations of part of these standards focused on particular scenarios and portions of the value chain, for instance content distribution to mobile phones.

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

² Open Mobile Alliance (OMA), <http://www.openmobilealliance.org>

copyright law is also a concern that has been criticised, for instance, by the Electronic Frontier Foundation [9].

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 Organisation³ promotes, there is a common legal base and fruitful efforts towards a greater level of copyright law worldwide harmonisation.

On the other hand, getting tied to the traditional access control approaches makes it very difficult to adapt DRM systems to new challenges and opportunities offered by Internet-wide content distribution. 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.

An alternative to existing DRM approaches is the Creative Commons initiative, whose objective is to promote content sharing and reuse through open licensing schemes. However, despite the success of Creative Commons licenses, which 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 formal representations of the available licensing terms. There are mainly six different Creative Commons licenses, all of them non-commercial, though an extension mechanism⁴ is being developed for alternative licensing schemes. In any case, though licenses are available as metadata for computers consumption, it is a simple representation intended to support simple services like content search. There are not mechanisms for advanced computerised support that enable an Internet-wide copyright-based alternative to DRM systems.

Ontologies and policy languages are one of the clear alternatives in order to formally represent content value chains and make it possible to implement copyright management systems on top of this formalisation. The importance of this alternative is exemplified by ISO/IEC MPEG recent interest in ontologies in the context of content value chains management. This interest is materialised in the establishment of the Adhoc Group on Requirements for Media Value Chain Ontology (MVCO)⁵. This group has already published a requirements document [10], which is the first step towards incorporating an ontology into the MPEG-21 framework.

In fact, there are already some ontologies for content management. OREL [11] is an ontology that tries to formalise part of MPEG-21 semantics. However, instead of providing a conceptualisation of the underlying semantics, it is based on a translation of part of MPEG-21 to a semantic representation language. Consequently, it inherits all the problems highlighted for DRM standardisation efforts.

A more complete formalisation is the RRD ontology for the expression of intellectual property entities and relations [12]. It is not based on copyright law but on a partial conceptualisation of the creation model, i.e. the different forms a creation can take along the value chain. However, this model is not sufficiently detailed as it just considers Works, Adaptations, Manifestations, Instances and Copies, as it is shown in Figure 1. There is no distinction between manifestations as objects or as events, e.g. a

³ WIPO, World Intellectual Property Organization, <http://www.wipo.int>

⁴ CCPlus, <http://wiki.creativecommons.org/Ccplus>

⁵ AhG on Requirements for MVCO reflector, <http://tech.groups.yahoo.com/group/mvco>

script or a performance, neither between instances as objects or as events, e.g. a recording or a broadcast.

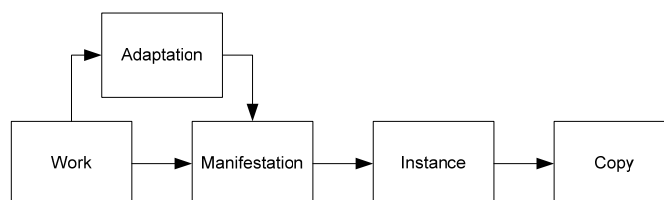


Figure 1. RRD Ontology Creation Model

Consequently, the kinds of value chains that can be modelled are restricted to simple and very direct ones, for instance a music work that is performed, recorded and then copies produced and distributed. However, it is not clear how this ontology can cope with complex scenarios like the one presented in Table 1.

Table 1. Broadcast of a serial adapted for a literary work scenario

<p>"A 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".</p>
--

It is clear that the literary work is a Work in RRD, the serial is an Adaptation and the script might be a Manifestation. However, the performance is not an Instance, the recording might be an Instance but the broadcast is not a Copy. In any case, someone recording the broadcast with his video might produce a Copy.

Consequently, it is clear that this ontology, though very concise and appropriate in order to model simple value chains, is not suitable for complex ones. The problem is that complex value chains seem to be the tendency as new ways to generate and distribute content proliferate, especially thanks to the facilities provided by the digital networks and the success of user generated content.

In fact, it can be observed that the RRD ontology creation model is inspired by the International Federation of Library Associations and Institutions (IFLA) FRBR⁶ Group 1 entities: Work, Expression, Manifestation and Item. FRBR entities show the same limitations if we want to develop a creation model that lays the foundations for a complete value chain model. FRBR Expressions can be both objects and events, consequently there is no distinction between scripts and performances and it is not possible to model that a particular performance depends on a particular script. Even broadcasts are considered expressions and there is no way to state that there is a recording bridging the performance and the broadcast. Finally, there is no way to represent derivative works, i.e. adaptation relations between works.

In any case, it is understandable that FRBR does not make these detailed distinctions as it was originally intended as a conceptual model for bibliographic records. However, other conceptual models intended for detailed content value chain modelling, as it has been shown for the RRD ontology, are based on the FRBR model and inherit this limitationsTable 1.

An alternative to the FRBR creation model is the OntologyX [13] ontology. As it is shown in Figure 2, it does distinguish between creations as objects, called

⁶ IFLA's Functional Requirements for Bibliographic Records, <http://www.frbr.org>

Manifestations, and as things in time, called Expressions. However, it does not further detail the different forms a creation can adopt along the value chain when as object or as event. Moreover, 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.

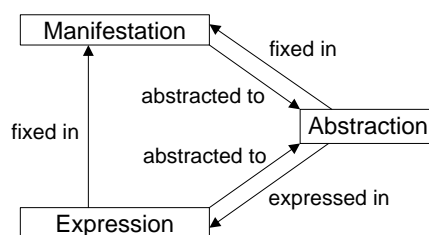


Figure 2. The OntologyX Creation Model

A part from existing ontologies for content value chains, there are other more generic initiatives in the area of Web rules and policies that can be also applied to this domain. We see all these initiatives as complementary to our work because they provide tools that allow modelling the dynamics of the business processes in the content domain. The more direct approach is to apply rules for business process integration by using rule markup languages [14]. A more elaborate approach is to also apply policies, which can be modelled on top of Web rule languages [15].

There are also generic policy languages based on ontologies like KAOs [16]. All these approaches are generic and do not provide specific constructs for content value chains modelling. However, policy ontologies like the KAOs one can be the basis on top of which an ontology for content can be developed. Moreover, a policy ontology presents some limitations that are best dealt with by rule languages, which on their turn show limitations that make an hybrid approach combining rules and ontologies the best choice [17]. Consequently, our idea is to combine the policy ontology approach with rule languages. Other ontologies that complement our contribution, as it is a solution that operates at the enterprise level, are enterprise modelling ontologies like REA [18] or UFO [19].

Finally, as we want to achieve the greatest level of interoperability, the best choice seems to take profit from existing Web ontologies [20] and Web rule languages [15]. This is the approach for the development of the Copyright Ontology⁷, detailed in the following section. The expressiveness and generality of the resulting conceptualisation allows coping with the shortcoming of existing approaches and, additionally, it can be used as an interoperability facilitator for the main DRM standards [21].

3. The Copyright Ontology

The copyright domain is quite complex so we face its conceptualisation in three phases that allow facing this process in an incremental way. First, the objective is the more primitive part, the Creation Model. Second, there is the model for the rights part, the

⁷ Copyright Ontology, <http://rhizomik.net/ontologies/copyrightonto>

Rights Model, and finally a model for the available actions, the Action Model, which is built on top of the two previous ones.

3.1. Creation Model

The Creation Model conceptualises the different forms a creation can take, which are first classified based on the common distinctions made by top ontologies [19,22,23]:

- **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 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 3, as it evolves from an abstract idea, i.e. a *Work*, towards something that can be consumed by end users, e.g. an *Instance*, a *Communication* or a *Performance*. Despite *Fixations* or *Manifestations* might also be consumed, e.g. a painting. This creation model constitutes a superset of the ones commented in the Related Work Section and consequently existing conceptualisations based on them might be mapped to the Copyright Ontology creation model.

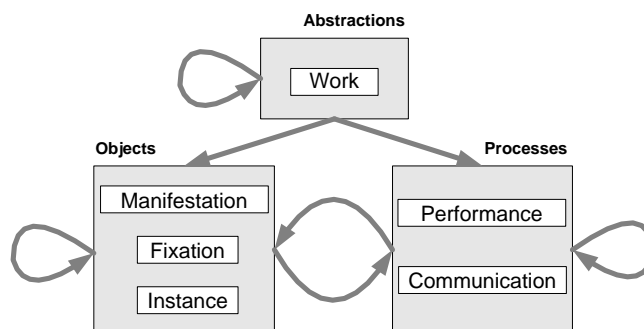


Figure 3. Creation model showing the different views on creation

3.2. Rights Model

Once the creation model is established, the foundations for value chain modelling are laid. However, it is necessary to define the dynamics that make creations move from one stage to the following one along their life cycle. These dynamics are shaped by the underlying copyright legal framework. It is modelled by the Rights Model, which follows the World Intellectual Property Organisation recommendations. It includes economic plus moral rights, as promoted by WIPO, and copyright related rights, see Figure 4.

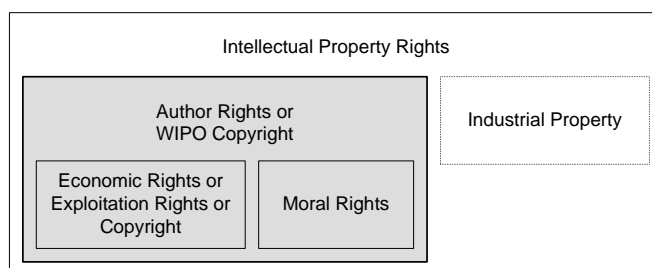


Figure 4. The Rights Model in the Copyright Ontology

The most relevant rights in the context of value chains are economic rights because 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)
- **Moral Rights:**
 - Attribution Right
 - Integrity Right
 - Disclosure Right
 - Withdrawal Right
- **Related Rights:**
 - Performers Rights
 - Phonograms Producers Rights
 - Broadcasters Rights

Copyright rights are used in different steps of the value chain. For instance, whenever an author generates a new creation, the author automatically gets all Economic and Moral Rights.

3.3. 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 5. Actions are the concrete things that value chain agents can perform and are regulated by the rights in the Rights Model. For the economic rights, these are the governed actions:

- **Reproduction Right:** to 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:** to fix something, *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*.

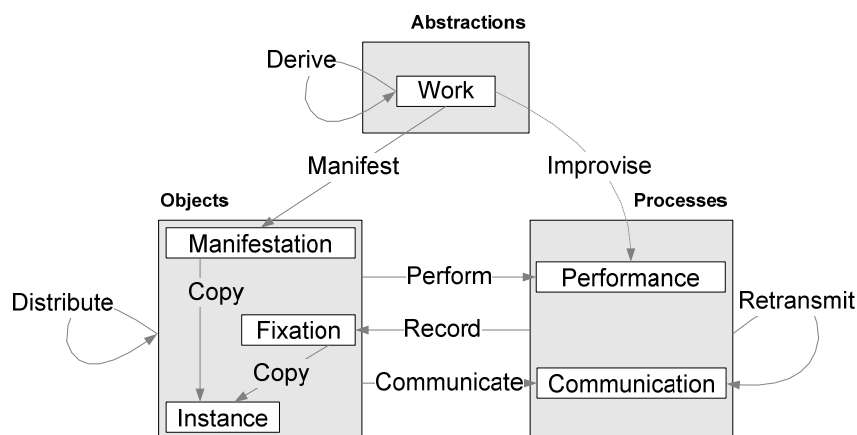


Figure 5. Relations between the Action and Creation Models

A part from the actions related to economic rights, there are also other actions with a great though indirect impact on value chains. They are related to one of the biggest criticisms against DRM, the special permissions that many copyright legal systems grant to end-users but are blocked by DRM systems. These permissions are commonly called fair use, fair dealing or user rights.

These privileged actions, normally restricted by copyright, may be done without the authorization of the copyright owner in circumstances specified by law. However, this does 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 levies constitute a significant part of the content market so this part of the value chain should not be ignored [24].

These are the main copyright exceptions:

- **Quotation Right:** to quote, a limited extent *Copy* action of a source protected work, which is clearly mentioned.

- **Education Right:** educational act, any *Copy*, *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 *copy* 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 *copy* act that produces a temporal reproduction.

All the previous actions are complemented with a set of relations that link them to their participants. This set is adopted from the linguistics field and it is based on case roles [25]. Case roles link the different parts of a sentence to its main component, the verb. In this case the verbs are the identified actions. The case roles in the Action Model are shown in Table 2.

The roles are categorised on four main types of roles, the columns. *Initiator* corresponds to a participant that determines the direction of the process from the beginning while *Goal* determines it for the end. *Resource* points to a participant that must be present at the beginning of the process, but does not actively control what happens, while *Essence* must be present at the end of the process. The roles also depend on the aspect of the verb they are related to, the rows. For instance, if the verb has a transfer aspect, something changes its “location”, or if it has a temporal aspect, how does it evolve along time.

Table 2. 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 previous pool of primitive actions and case roles, together with the restrictions that relate each action to the relevant case roles [26], allows modelling the concrete events that shape value chains in the copyright domain. For instance, Figure 6 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, a manifestation that is performed by some actors, e.g. Gerard Depardieu, and recorded into a motion picture. This motion picture is finally broadcasted to consumers who can tune the resulting communication and even record it producing a fixation.

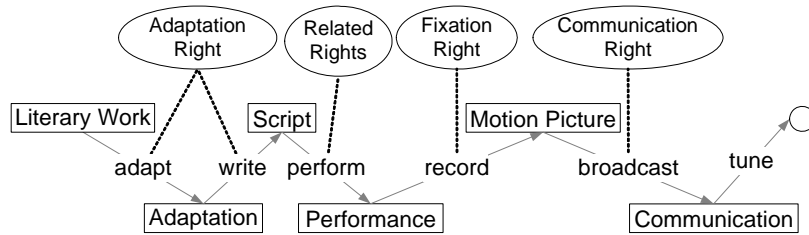


Figure 6. Literary works adapted to serials value chain

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 with the corresponding action and associated participants linked 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 [27], the modal operators that make it possible to model things like licenses and contracts [28].

3.3.1. Copyright Licenses Modelling

Copyright provides a legal framework that governs creations life cycle 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 interactions 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, the modal part.

In order to produce a more homogeneous and usable conceptualisation, we have incorporated these concepts in the ontology using the terms that appear in licenses when expressed using natural language, i.e. using the corresponding actions and case roles (e.g. Agree for the verb "to agree" used in licenses in order to model a permission).

However, prior to incorporating modal operators, event patterns should be introduced as the way to state what is obliged, permitted or prohibited by a license; they are then 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 7 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".

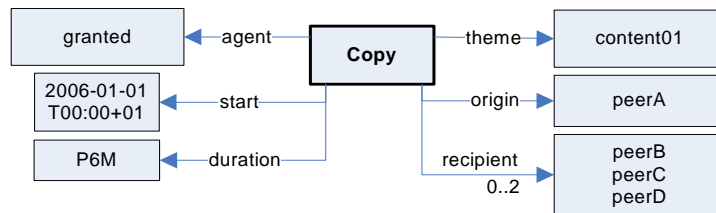


Figure 7. Pattern for a copy action in a P2P scenario

Then, there are the terms to state the modality of these event patterns in copyright licenses. The Copyright Ontology does so 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.

Permissions are captured by a new action, *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 7, in order to authorise the pattern that it models, an agreement like the one shown in Figure 8 can be modelled. The agreement between “granter” and “granted” in the upper part authorises the pattern pointed by the *theme* case role, the previous P2P copy pattern at the centre of the figure.

Obligations are captured 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 8 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. There is also a cardinality constraint that clarifies that just one transfer is required.

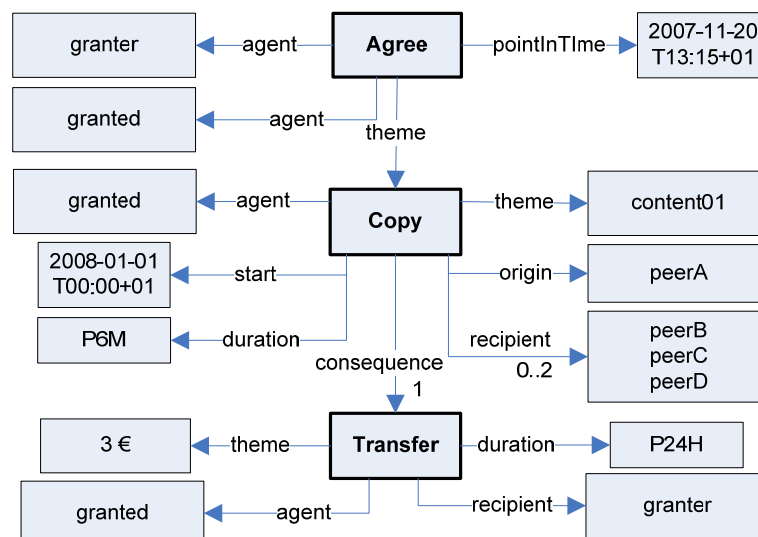


Figure 8. Agreement that permits the P2P copy pattern whose consequence is an economic obligation

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*.

Finally, conditions 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 them. It is applied to the pattern that is conditioned and it links to the pattern that establishes the condition. 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 be conditioned by a *Transfer* one that requires that the “granted” agent makes a 1 Euro prepayment to the “granter” agent before the former can exercise the permitted P2P *Copy* action.

This completes the static part of the Copyright Ontology. For a full account of the ontology conceptualisation and additional entities reused from other ontologies in order to make it possible to model related aspects like the kinds of involved parties or mereological relations, please refer to [26].

4. Modelling Value Chains Dynamics

The previous ontology concepts and relations model the static part of value chains, being the more complex constructs the events that capture the static “pictures” that shape, step by step, value chains. Therefore, it is necessary to incorporate rules and reasoning mechanisms that make possible to capture the domain dynamics that make value chains move.

In order to do that, we have followed a hybrid approach that combines ontology reasoning mechanisms and rules [17]. First, there are the ontology reasoning mechanisms based on Description Logics that implement license checking and authorise the uses granted by licenses, which are detailed in Section 4.1. For complex policies and for modelling the rest of the dynamic aspects, rules are applied as detailed in Section 4.2.

4.1. Ontology-based License Checking

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 ontology has been implemented using the Description Logic (DL) variant of the Web Ontology Language (OWL-DL), which is constrained in order to be managed by DL reasoners.

The implementation approach, similar to the KAoS one [16], it profits from DL reasoners in order to implement part of the domain dynamics. They 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.

Patterns are implemented as OWL classes 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. The main restrictions in OWL 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, pertain to the time range [2008/01/01–2008/06/30]. In order to support the later, custom datatypes reasoning is required [29].
- **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 action patterns. For instance, Figure 9 shows the conceptual model for a license that combines commercial and open access terms.

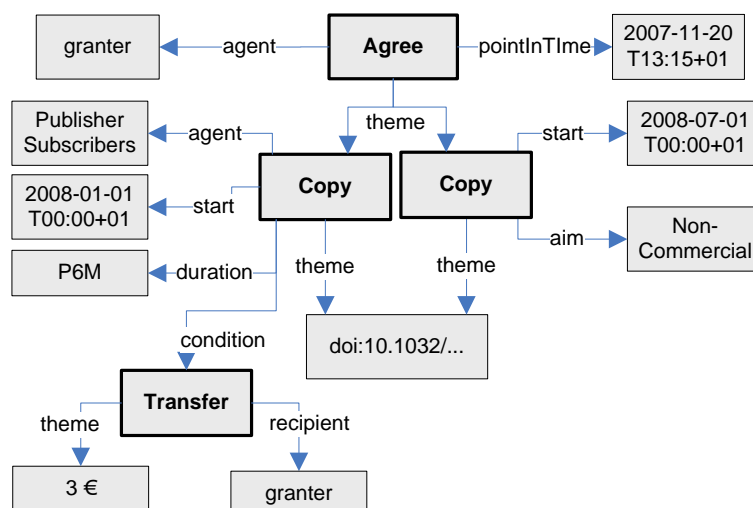


Figure 9. Agreement on a copy action under commercial and open access terms

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 10 shows the pattern build up from the combination of such kind of restrictions for the example presented in Figure 9. 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.

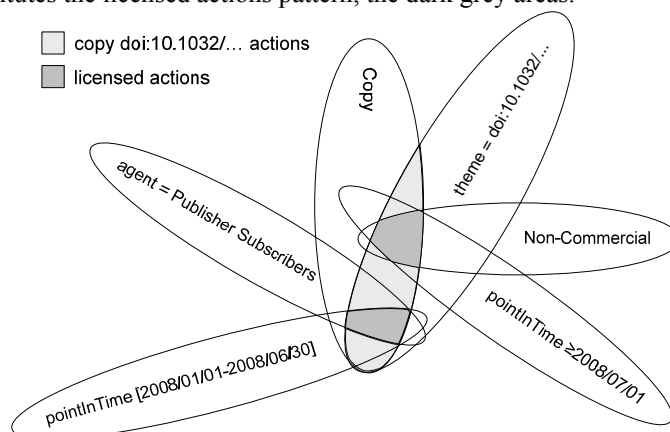


Figure 10. Building an action pattern as an intersection of restrictions

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 2008-07-01. Table 3 shows the DL notation for the class definition that models the commercial copy pattern.

Table 3. OWL-DL Class for the commercial copy action pattern

Pattern \equiv Copy \sqcap	(1)
\forall pointInTime. $\geq 2008-01-01T00:00:00, \leq 2008-06-30T23:59:59$ \sqcap	(2)
\forall agent. PublisherSubscribers $\sqcap (\geq 1$ agent) \sqcap	(3)
\exists theme. {urn:doi.10.1032/...} $\sqcap (\leq 1$ theme)	(4)

Each intersected restriction reduces the initial set of actions, which corresponds to all the *Copy* actions. First, (1) 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, (2) and (3), 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 later 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*.

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 [28].

Despite the versatility of the DL based approach to license checking, as it has been shown in the previous lines, there is an underlying algorithm that guides what to check using the DL reasoner and how to interpret its results in the context of the permissions, prohibitions and obligations that operate on event patterns. As it is detailed in the next section, this part can be implemented using rules or programmatically for restricted scenarios.

4.2. Rule-based Modelling

There are many dynamic aspects of content value chains modelling that cannot be captured using DL reasoning mechanism. For this part, we have applied a rule-based approach. This part has not been fully implemented because it tends to be too scenario specific. In fact, for many scenarios where the Copyright Ontology has been applied, these part has been implemented using alternatives like procedural code or SPARQL queries [30] due to its limited scope and complexity.

If the scenario just considers a part of the value chain, especially if it is the part nearer to consumers, the license checking dynamics captured by the OWL-DL implementation reduce the remaining behaviours to be implemented to the interpretation of the modal part that guides DL reasoning. This has make it easier and more efficient to implement them procedurally for scenarios like a copyright-aware streaming server or P2P controlled metadata diffusion [31]. In these cases, the range of actions involved is mainly reduced to *Access* or *Copy* and the dynamic part to licenses checking. Moreover, it is also possible to use SPARQL queries in order to implement patterns requiring variables and closed world assumptions [32].

However, for more complex scenarios set of rules have been developed. The first case is to model the creative process, how authors acquire rights just as a result of manifesting a new work or by deriving one from an existing work. For instance, Table 4 presents the conceptualisation of the rule that, given a manifest event when the creator materialises a work, assigns the moral and economic rights on that creation to the creator. There are also rules for other creative processes but they have just been conceptualised [26], there is not any implementation based on Web rule languages yet.

Table 4. The Manifest Rule that assigns rights to creators

$ \begin{aligned} & (\forall v:\text{Manifest}) \\ & ((\exists mr:\text{MoralRights})(\exists er:\text{EconomicRights}) \\ & ((\forall p:\text{Person})(\forall m':\text{Manifestation})(\forall w:\text{Work})(\forall m:\text{Manifestation})(\forall t:\text{Time}) \\ & (\neg\text{isManifestationOf}(m',w) \wedge \text{agent}(v,p) \wedge \text{theme}(v,w) \wedge \text{result}(v,m) \wedge \\ & \text{pointInTime}(v,t) \wedge \text{location}(v,l) \\ & \rightarrow \\ & \text{agent}(mr,p) \wedge \text{agent}(er,p) \wedge \text{essence}(mr,w) \wedge \text{essence}(er,m) \wedge \\ & \text{start}(mr,t) \wedge \text{start}(er,t) \wedge \text{location}(mr,l) \wedge \text{location}(er,l) \wedge \text{isManifestationOf}(m,w))) \end{aligned} $
--

A complete set of rules have been developed and implemented as an extension⁸ to the Copyright Ontology for a pair of scenarios related with user rights. These scenarios are private copy and quote [33]. The following section provides details about how private copy has been implemented using rules. This is a particularly important scenario to take into account as private copy justifies an important revenue channel for authors through the levies collected by collective societies. Many DRM systems block this user right and, consequently, this part of the value chain.

4.2.1. Private Copy

Private copies are all copy actions performed by any person on any content instance that have as a result a replica instance whose aim is private consumption. The corresponding conceptualisation in Copyright Ontology terms is shown in the upper part of Figure 11. The *private* term is modelled as the value of the *aim* case role.

In order to detail the private consumption part, there is also an agreement on any use of the resulting instance by any agent directly related to the person that produced the private copy, which is shown in the bottom part of Figure 11. Any other uses by persons not directly related are not allowed if they are not explicitly granted.

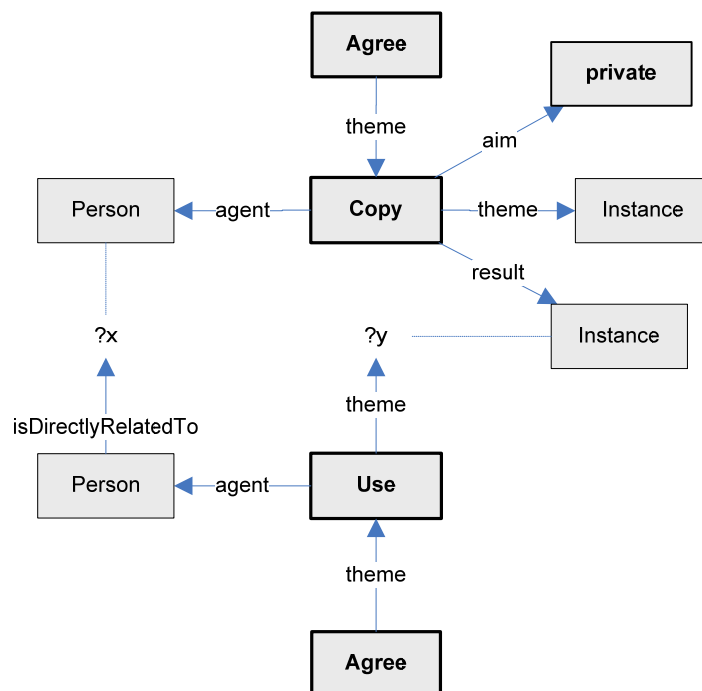


Figure 11. Model for the “Private Copy Right”

Two things must be highlighted about the second agreement. First, the *isDirectlyRelatedTo* relation is used as a way to model direct relations among people and tries to capture the private essence, e.g. family, friends, etc. Second, the variables

⁸ Private Copy Ontology, <http://rhizomik.net/ontologies/2007/10/privatecopy.owl>

“?x” and “?y” are used in order to state that the relation must hold to the same person that performed the copy and to the same resulting instance.

Named variables are not available in OWL-DL but they are captured by Semantic Web rules. In this case, a rule is in charge of generating the *Use* pattern and the second agreement using the *Copy* pattern as input. Whenever an action is classified as an instance of the *Copy* pattern, the rule is triggered and it asserts the agreement with the *Use* pattern class for the concrete person and instance.

A first implementation of the rule conceptualised in Figure 11 has been implemented in Jena using a Jena rule⁹. It is based on a *Copy* pattern that, when it is matched, triggers the rule that asserts the *Use* pattern that authorises the use of the private copy to people directly related to the private copy creator. Consequently, from this moment, any use that is classified into the class corresponding to the pattern will be granted.

To conclude, it is important to note that the ontology focuses on modelling the private copy scenario, not on enforcing it. Existing enforcement frameworks, like those based on MPEG-21 might be used to this end. For instance, it has been already proposed to use the copyright ontology [34] as one of the components of the Chillout platform, an open source implementation of an interoperable DRM platform¹⁰. The copyright ontology and the reasoning capabilities based on it might be integrated in the License Provider Device, a central component of the Chillout architecture responsible for managing licenses and requests from end-user devices.

Moreover, the ontology might be also used to support rights management systems based on accountability instead of on enforcement, e.g. those based on watermarking measures [35].

From the point of view of the ontology, the previous actions are just annotated and they are legal as long as there is not any counter evidence. For instance, it could be discovered that the person using the copy is in fact not directly related to the private copy maker. Consequently, the DL reasoner would detect an inconsistency because the class of persons directly related to another one is disjoint with the class of persons that are not directly related to this same person.

On the other hand, if DRM enforcement is in place, the private use part of the model, and more specifically the *isDirectlyRelatedTo* relation, can be enforced, for instance using some sort of device domains and encryption measures as the ones specified in the context of mobile devices by OMA¹¹.

5. Conclusions and Future Work

The Copyright Ontology is an ontology based on the copyright legal framework that is capable of modelling complex value chains as compared to traditional DRM languages, open initiatives like Creative Commons or other existing digital rights ontologies.

First of all, it provides a complete creation model that takes into account the different forms that creations can take along the content value chain. Second, there is the rights model that captures concepts from the underlying legal framework and gives

⁹ Jena Inference Support, <http://jena.sourceforge.net/inference>

¹⁰ Chillout, <http://chillout.dmpf.org>

¹¹ DRM Architecture. Open Mobile Alliance, Approved Version 2.0, 2006.
http://www.openmobilealliance.org/release_program/drm_v2_0.html

shape to the third component, the action model. The latter includes the concrete actions governed by copyright that can be performed by value chain participants, which makes it easier copy copyright managers to write and understand licenses based on these building blocks.

Moreover, there are special actions and case roles that allow modelling copyright licenses by relating actions to their participants and to modal operators for permissions, prohibitions and obligations.

All the previous concepts and relations constitute the static part of content value chains model. In order to capture the dynamic part, a hybrid approach combining ontology-based and rule-based mechanisms has been used. The Copyright Ontology has been implemented using OWL-DL in such a way that DL reasoners can be used in order to perform license checking and authorise uses granted by licenses.

However, for complex aspects of value chains, rules have been used. Currently, there is not a complete implementation of the dynamic part of value chains using rule languages. Most of the scenarios where the Copyright Ontology has been applied have limited scope and it has been possible to implement them using a programmatic approach. However, the creative process by which rights are assigned to creators has been conceptualised using rules due to its complexity.

The scenarios that have been completely implemented using rules are related to user rights, concretely private copy and quote. A part from the importance of private copy in content value chains, the complexity of these scenarios, usually ignored by DRM systems and other digital rights ontologies, serve as a showcase of the possibilities of the Copyright Ontology when combined with rule languages.

The future work concentrates now on completing the rule-based model for the dynamic aspects. The main objective is to implement the aspects of license checking not covered by the ontology-based approach, which are currently implemented procedurally or using SPARQL queries. A rule-based implementation of this part will make the whole system more flexible and easier to apply to new scenarios.

Finally, we would like to take profit from the natural language processing roots of the action model to facilitate formalizing licenses from their natural language expression, an easier way to express licenses for digital rights managers.

Appendix A. Analysis of MPEG-21 Media Value Chain Ontology Requirements

The Copyright Ontology constitutes a choice as the ontology for media value chains modelling in the context of ISO/IEC MPEG-21 Media Value Chain Ontology standardisation process. We have been participating in the requirements gathering phase and, below, there is a list of the preliminary requirements [10] and a short description about how the Copyright Ontology satisfies them:

- **Common core semantics:** the ontology shall define a core set of terms common to any copyright management system. The Copyright Ontology does so taking into account the underlying copyright legal framework that shapes the corresponding content value chains.
- **Consistency:** as a formal model, consistency can be checked by logic reasoning means and guaranteed.
- **Concurrence between Machine and Human Readable Semantics:** the ontology is accompanied by a human readable documentation. Moreover, we

are currently exploring ways to render Copyright Ontology expressions using Controlled Natural Languages [36].

- **Machine readable:** the ontology is implemented using Semantic Web tools that make it machine readable.
- **Core extensions:** the ontology is easily extendable using Web Ontology Language constructs. For instance, there is an extension for private copy.
- **Scope of value chain participants:** the whole value chain and all the potential participants should be considered. The copyright ontology provides a detailed creation model and set of actions, governed by copyright, that allow modelling complex value chains. The participants are linked through a set of generic case roles that facilitate extensibility.
- **Core model and relations:** the ontology contributes a core model based on the creation, rights and action models. This core model is enriched with dynamic aspects captured by ontology-based and rule-based mechanisms.
- **Compatibility with existing IP representation:** existing intellectual property identifiers (e.g. ISWC, ISRC, etc.) can be reused in Copyright Ontology expressions as URIs.
- **Authorisation, attribution and licensing:** licenses or uses modelled using the Copyright Ontology can be used for license checking, use authorisation and audit. The enforcement mechanisms are provided by other parts of the MPEG-21 framework.
- **Common digital and non digital governance:** the ontology is based on the underlying legal framework that includes both digital and non digital content.
- **Copyright exceptions:** the ontology gives support to copyright exceptions like private copy or quote.
- **REL integration and support:** The MPEG-21 Rights Expression Language can be integrated with the ontology through the REL Ontology¹² [21].
- **Unfettered access:** the ontology is public and licensed under non-commercial, attribution required and share alike terms in order to make access to the ontology as easy and inexpensive as possible.
- **Extension of MPEG-21 standards:** the ontology complements MPEG-21 standards as it just deals with content value chains modelling and can reuse many MPEG-21 parts for its operation, for instance enforcement or event reporting mechanisms.
- **Social tagging:** user generated metadata can be associated with any concept or relation of the Creation Ontology, or with any expression based on it. The use of Semantic Web representation means make it possible to profit from URIs in order to do so.

References

- [1] Lessig, L.: “The Future of Ideas: The Fate of the Commons in a Connected World”; Vintage (2002)
- [2] Becker, E., Buhse, W., Günnewig, D., Rump, N. (eds.): Digital Rights Management: Technological, Economic, Legal and Political Aspects. Springer, 2003.

¹² Semantic DRM System, <http://rhizomik.net/semdrms>

- [3] Chiariglione, L.: "Ad hoc groups established at the 83rd MPEG meeting, AHG on Requirements for Media Value Chain Ontology". Mandate 9657. http://www.chiariglione.org/mpeg/meetings/antalya08/antalya_ahg.htm
- [4] de Walle, R.V., Burnett, I.: "The MPEG-21 Book"; John Wiley & Sons, UK, 2005.
- [5] Rosenblatt, B.: "2005 Year in Review: DRM Standards"; DRM Watch, January 2, 2006 <http://www.drmwatch.com/standards/article.php/3574511>
- [6] 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.
- [7] 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, pp. 137-146.
- [8] Iannella, R.: "Open Digital Rights Language (ODRL), Version 1.1". World Wide Web Consortium Note, 2002. <http://www.w3.org/TR/odrl>
- [9] Doctorow, C.: "Critique of NAVSHP (FP6) DRM Requirements Report"; Electronic Frontier Foundation (2005) <http://www.eff.org/IP/DRM/NAVSHP>
- [10] "Requirements for a Media Value Chain Ontology". ISO/IEC JTC 1/SC 29/WG 11, Approved Public Document N9658. January 2008, Antalya, Turkey. http://www.chiariglione.org/mpeg/working_documents/explorations/mvco/MVCO-Reqs.zip
- [11] Qu, Y., Zhang, X., Li, H.: "OREL: an ontology-based rights expression language". In Proceedings of the 13th international World Wide Web Conference. New York: ACM Press, 2004, pp. 324-325.
- [12] Rodríguez, V., Gauvin, M., Delgado, J.: "An Ontology for the Expression of Intellectual Property Entities and Relations". In M. I. Yagüe & E. Fernández-Medina (eds.) Security in Information Systems, Proceedings of the 5th International Workshop on Security in Information Systems, WOSIS 2007. Madeira, PT: INSTICC Press, 2007, pp. 196-203.
- [13] Pease, A., Rust, G.: "Formal Ontology for Media Rights Transactions". In R. García (ed.) "Semantic Web for Business: Cases and Applications". IGI Global, in press, 2008.
- [14] Milanovic, M., Kaviani, N., Gasevic, D., Giurca, A., Wagner, G., Devedzic, V., Hatala, M.: "Business Process Integration by Using General Rule Markup Language". In Proceedings of the 11th IEEE Enterprise Distributed Object Computing Conference, EDOC'07. IEEE Computer Society, 2007, pp. 353-364.
- [15] Kaviani, N., Gasevic, D., Hatala, M., Wagner, G.: "Web Rule Languages to Carry Policies". In Proceedings of the 8th IEEE Workshop on Policies for Distributed Systems and Networks. IEEE Computer Society, 2007, pp. 188-192.
- [16] Uszok, A., et al.: "KAoS policy management for semantic web services"; IEEE Intelligent Systems, 19, 4 (2004), 32-41.
- [17] Toninelli, A., Bradshaw, J., Kagal, L., Montanari, R.: "Rule-based and Ontology-based Policies: Toward a Hybrid Approach to Control Agents in Pervasive Environments". In Proceedings of the Semantic Web and Policy Workshop, 2005.
- [18] Gailly, F., Poels, G.: "Towards an Operational REA business ontology". In R. García (ed.) "Semantic Web for Business: Cases and Applications". IGI Global, in press, 2008
- [19] Guizzardi, G., Wagner, G.: "Some applications of a unified foundational ontology in business modelling". In P. Green and M. Rosemann (eds.) Business Systems Analysis with Ontologies. Idea Group Publishing, Hershey, 2005, pp. 345-367.
- [20] McGuinness, D.L., van Harmelen, F.: "OWL Web Ontology Language Overview"; W3C Recommendation 10 February 2004. <http://www.w3.org/TR/owl-features>
- [21] García, R., Gil, R., Delgado, J.: "A web ontologies framework for digital rights management"; Journal of Artificial Intelligence and Law, 15, 2 (2007), 137-154.
- [22] 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, pp. 2-9.
- [23] 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
- [24] Graber, C. B., Govoni, C., Girsberge, M., Nenova, M. (eds.): "Digital Rights Management: The End of Collecting Societies? Huntington, NY: Juris Publishing, 2005.
- [25] Sowa, J.F.: "Knowledge Representation. Logical, philosophical and computational foundations"; Brooks Cole Publishing Co., 2000.
- [26] García, R.: "A Semantic Web Approach to Digital Rights Management". PhD Thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2005. <http://rhizomik.net/~roberto/thesis>
- [27] Lee, R.M.: "A Logic Model for Electronic Contracting"; Decision support systems, 4 (1988), 27-44.

- [28] García, R., Gil, R.: "A Web Ontology for Copyright Contracts Management". *Journal of Electronic Commerce*, (2008), in press.
- [29] Sirin, E., Parsia, B., Grau, B., Kalyanpur, A., & Katz, Y.: "Pellet: A practical OWL-DL reasoner". *Journal of Web Semantics*, 5, 2, (2007), 51-53.
- [30] García, R., Pariente, T.: "Copyright Management for the LUISA Semantic Learning Content Management System". 1st World Summit on the Knowledge Society, WSKS'08. *Lecture Notes in Artificial Intelligence*, Vol. 5288, Springer, 2008, pp. 253-262.
- [31] García, R., Tummarello, G.: "Semantic Digital Rights Management for Controlled P2P RDF Metadata Diffusion". In 2nd International Semantic Web Policy Workshop, SWPW'06. *CEUR Workshop Proceedings*, 207, 2006, pp. 102-109.
- [32] Gandon, F.: "SPARQL in a nutshell". Slideshare, 2008.
http://www.slideshare.net/fabien_gandon/sparql-in-a-nutshell/
- [33] Springer, M., García, R.: "Promoting Music Sampling by Semantic Web-Enhanced DRM Tools". In R. Grimm, B. Hass (eds.), *Virtual Goods: Technology, Economy, and Legal Aspects*. Nova Science Publishers, 2008, in press.
- [34] García, R., Springer, M.: "Formalising TRUs with the Copyright Ontology". Document No. 1010, 15th General Assembly, The Digital Media Project, August 18th, 2007.
<http://www.dmpf.org/open/dmp1010.doc>
- [35] Kravets, D.: "DRM Is Dead, But Watermarks Rise From Its Ashes". *Wired Magazine*, January 2008.
- [36] 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*. Springer, 2004, pp. 711-717.