Formalising a Model for Digital Rights Clearance

Claudio Prandoni[1], Marlis Valentini[1], and Martin Doerr[2]

[1] Metaware S.p.A., Via Filippo Turati 43/45,
56125 Pisa, Italy
{c.prandoni,m.valentini}@metaware.it
[2] Institute of Computer Science, FORTH,
Vassilika Vouton, 71110 Heraklion, Crete, Greece
martin@ics.forth.gr

Abstract. Due to the increasing complexity and world-wide distribution of digital objects, identification and enforcement of digital rights have become too complex to be carried out manually. It is necessary to take into account the case-specific applicable laws, the complete creation history of a work and the existing licenses. However, no formal generic model has been presented so far integrating these aspects. This paper presents an innovative domain ontology of the Intellectual Property Rights. It distinguishes four levels of abstraction or control: (1) the legal framework, (2) the individual rights people hold, (3) the individual usage agreements right holders and others may issue, and (4) the particular actions that are restricted by IPR regulations or bring particular rights into existence. The ontology has the potential to enable wide semantic interoperability of digital repositories for identifying existing rights on digital objects and tracing the impact of particular actions on rights and regulations.

Keywords: Intellectual Property Rights, Digital Provenance, Digital Rights, Information Integration, Interoperability, Ontology, Domain Models, Ontology Engineering.

1 Introduction

Intellectual Property Rights (IPRs) are exclusive rights that the law grants to the authors of intellectual creations. They are divided in Industrial Property Rights and Copyright. Industrial Property Rights covers inventions (patents), industrial designs, trademarks and indications of source, while Copyright protects literary, artistic and scientific works [1]. Related Rights were introduced later to complement Copyright and to protect artistic performances, sound productions, databases, software and folklore.

This paper presents a domain ontology, hereinafter called the "Digital Rights Ontology" (DRO), that aims at providing a comprehensive conceptualisation of the basic relationships and entities in the Intellectual Property Rights domain, focusing in more detail on Copyright and Related Rights. This ontology has been developed in the framework of the European-funded IST project CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval - IST-2005-2.5.10) in order to support the definition and the tracing of digital rights in the context of long term access and preservation of digital archive holdings.

M. Agosti et al. (Eds.): ECDL 2009, LNCS 5714, pp. 327–338, 2009.
© Springer-Verlag Berlin Heidelberg 2009

2 Application Scenarios

The Digital Rights Ontology presented in this paper aims at modelling within one single framework the most relevant aspects required in digital rights applications. These can be divided at a general level in two main categories:

- identification and reporting of rights;
- evaluation of rights and consequences of actions.

An archive must be able to determine Intellectual Property Rights because they impose legal limitations to the actions on the archived material. Final consumers are not allowed, by law, to freely use rights-protected creations unless they have the proper permissions. The same is true for the archival institution; unless it has obtained the complete rights ownership by means of a transfer of rights, it needs an explicit authorisation to hold any copy, to perform any kind of modification and to distribute it. If it does not possess these permissions, the right holder might have, in the worst case, the right to ask for compensation payments and for deleting all copies, which would mean not only the loss of an archive holding, but also the loss of all the effort spent to curate and preserve it through the years.

In order to identify all existing rights on digital objects, the *Digital Provenance* is the first aspect that must be analysed. Digital Provenance documents the actions that form the creation history of a digital work and its constituent parts; providing, among others, information on who participated in the creation and what their contribution was. A complication arises when there is a relationship with other existing creations; for instance, the incorporation of parts of expressions in other expressions [4], such as the use of images, sound tracks, etc., in multimedia productions. This requires the ability to trace inclusion and derivation chains in the production.

Further elements that must be considered are the country under which a creative work is protected and the happening of particular events, such as transfers of rights, death of an author and consequent copyright expiration after a given time, or voluntary release of a work under Public Domain.

The second important aspect in the digital rights clearance workflow is the evaluation of rights and the consequences of actions. In this case an analysis of existing licenses and applicable laws is required. This analysis then leads to the compilation of a set of rules, which can be modelled as patterns of constrained activities against which to check the specific individual actions performed by a particular user on a particular work. In this way, implementing a rights enforcing algorithm means to check if there is at least one pattern against which the single individual situation matches. If a match is found then the action is allowed, otherwise it is assumed to be prohibited. Detecting allowances rather than prohibitions guarantees that no illegal action is regarded to be permitted by the system, even if the information the decision is based on may be incomplete; however, an allowed case may be missed by the system. The precision of the result depends on how much information is available on the digital provenance and the licensed rights. In case of complicated licenses, the evaluation system may indicate potential allowances to the user that it cannot evaluate automatically.

3 Existing Rights Models

In Digital Rights Management (DRM) systems, technical enforcement measures are often used in conjunction with machine-understandable Rights Expression Languages (RELs). The latter are used to express license terms and conditions that can be checked automatically. Some standards already exist, such as MPEG-21REL and ODRL (Open Digital Rights Language). MPEG-21REL incorporates XrML and is probably the most widely adopted specification for commercial DRM; while ODRL was created as an international effort to provide an open standard for the DRM sector. Both come with their own rights model [5][6]. However, the types of rights that they conceptualise are only the permissions obtained by the consumers. Ownership rights, i.e. the exclusive rights held by the authors of the creative works, are not modelled.

Besides these standards, there are the Creative Commons licensing models [7] and the associated ccREL [8][9] language. CC licenses, likewise other Copyleft [10] licensing forms, aim at releasing most of the limitations that derive from Copyright. ccREL allows to specify permissions and related constraints, but not ownership rights.

In addition to RELs, other sources of information have also been consulted. The work carried out by R. García, in particular the analysis and conceptualisation reported in [11], and implemented in the ontologies IPROnto [1] and RightsOnto [12], has been used as a starting point. The driver of these ontologies was the management of IPR licenses in e-commerce multimedia applications. They model the core concepts for checking and retrieving usage permissions through DL reasoning and semantic queries, providing an alternative to the previous syntactic approaches to build interoperability between different RELs.

It is worthwhile to mention the <indecs> framework [13] as the precursor of many of the current existing rights metadata formats, both in the academic research and in the commercial field (MPEG21 RDD). This project has recognised that mechanisms to transform metadata into representations of events of Digital Provenance appear to provide the most powerful approach to interoperability. The event-based model in <indecs> does, however, only allow the user to trace rights, as was required in e-commerce systems, and to retrieve events, but it does not model how rights effect events.

The DRO was developed independently from any particular service implementation and, with respect to the aforementioned right models, it aims to provide a more comprehensive conceptualisation of rights. It includes ownership rights as well as permissions, and all the relevant entities related to provenance, content of legislation and content of licenses which influence the status of rights.

4 The Digital Rights Ontology

4.1 Modelling Approach

The development of the DRO builds on some existing standard or well-established core ontologies, such as CIDOC CRM [14][15][16] and FRBRoo [15][17][18][19].

The CIDOC (Documentation Committee of the International Council of Museums) Conceptual Reference Model provides definitions and a formal structure for

modelling concepts and relationships used in cultural heritage documentation, and became an international standard, ISO 21127:2006. It is an event-centric model of the material history of people, their interactions and the objects they deal with. It has been in development since 1996 by abstracting over an empirical base of hundreds of database schemata and data structures from different cultural-historical domains. It already integrates the basic ideas of the <indecs> framework [14] and the subsequent ABC Harmony Model, which also describes the history of digital artefacts [20].

FRBRoo is an ontological interpretation created in 2003-2008 from FRBR (Functional Requirements for Bibliographic Records), the major expression of library conceptualisations created by IFLA (International Federation of Library Associations and Institutions) in the years 1992-1998. It identifies bibliographic entities (such as works, persons, events, places, etc.), their attributes and the relationships between them. It is formulated as a specialisation of ISO21127 and it is the result of careful harmonisation of the CIDOC and FRBR models. A major innovation of FRBRoo is the modelling of the mechanisms of *incorporation* of works, such as the creation of a song incorporating lyrics from an existing poem, or the use of a reproduction of a painting on a book cover.

The extension with certain concepts and terminology of the intellectual property rights domain relies on existing IPR-specific works. Indeed, several concepts have been adopted from the ontologies IPROnto and CopyrightOnto or reinterpreted in order to conform with the ontological rigor applied to the DRO. However, taxonomy of the adaptation is substantially different. In fact, the IPR-relevant actions were modelled in IPROnto as subclasses of the rights that govern that class of action. For instance, the *CommunicationRight* has *Communicate, Broadcast, Retransmit* and *MakeAvailable* as subclasses. As explained in [11], this ontologically wrong connection between rights and actions allowed the reduction of checking if an action is authorised, to just check if the right class subsumes the action class. However, it fails to represent correctly any other relevant events and to integrate other models. Instead, in the approach followed for the definition of the DRO, the checking of permissions does not consist in performing class subsumption, but matching a set of ontology instances (the circumstance of an intended action) against another set of ontology instances (the pattern that describes a certain permission).

In addition, some standard Rights Expression Languages, like MPEG21-REL [5], ODRL [6] and Creative Commons [8] have been used as a guide in the definition of the license-specific concepts and terminology. Finally, the documentation from WIPO (World Intellectual Property Organisation) has been used for consultation.

The extension and harmonisation of existing ontologies outlined above makes it possible not only to focus on the modelling of rights, but also to be aligned with the view of the curators (mainly library and museum archivists) and the view of the legislators (for the notion of ontological adequacy see [21]). This will allow, for instance, to simplify the mediation between different information sources, as long as they commit to a common ontology.

4.2 Rights as a Pattern of Constrained Activities

The core idea that characterises the Digital Rights Ontology is the distinction of 4 levels of abstraction or control, as indicated in Fig. 1:

- *legal framework*, consisting in the type of rights that are valid in a given country at a given time;
- *individual rights* owned by the right holders (also called *ownership rights*), e.g. Peter, creator of a digital work, owns the right to make a copy of it;
- *individual usage agreements* between right holders and other people, e.g. Peter issues to Martin a license which allows him to make copies of Peter's work under certain conditions;
- *particular actions*, which may fall under IPR-related regulations, e.g. Martin makes a copy of Peter's work. They may be allowed, disallowed, or allowed under particular conditions.

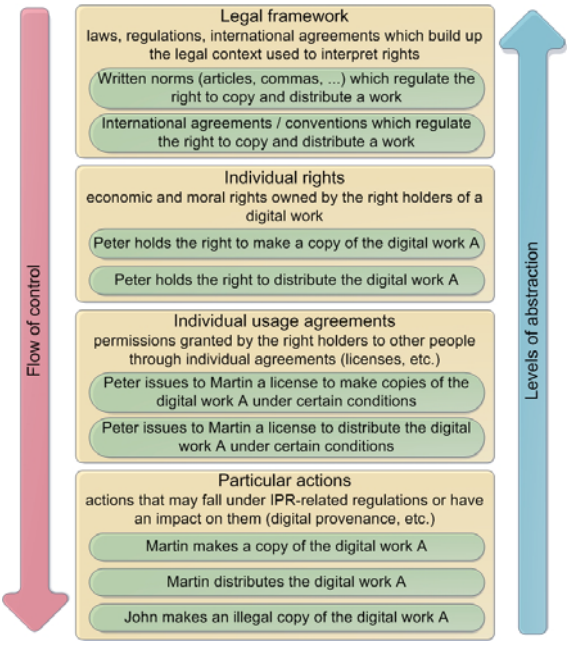

Fig. 1. Levels of abstraction or control

These levels can be also interpreted in a top-down manner as a flow of control: the legal framework documents and regulates the individual rights owned by the right holders; the latter regulate the individual permissions that right holders are allowed to grant to other people; and finally the individual permissions granted by the right holders, if any, regulate the particular actions that a user can legally perform on a digital work, overriding general prohibitions.

The underlying idea at the base of the DRO is that all the three upper levels: laws, rights ownership and licenses, are modelled as patterns of constraints for activities that determine if an activity is permitted or not, or if subsequent activities of particular patterns are foreseen. This means that each specific individual action is related to the agreements, rights and laws that regulate it. Martin can make a copy of a digital work

because he either signed an individual agreement with the creator of the work, he is the right holder of the work, or there is a particular law which grants him the right to make such a copy. This is a pattern matching process that should be applied at three different levels and can be implemented by suitable procedural code.

This way to make the nature of rights explicit allows for formalising one of the key processes needed in the domain of rights, namely the checking of permissions to perform a certain activity. This process can be described as a matching procedure, where the permissions (patterns) are matched against a given situation. This is in fact what happens when permissions are verified, either if they are given by the right holders through licenses or if they are given by the law in form of "user rights". A concrete example of how this mechanism works is described in paragraph 5.2.

4.3 Description of the Core Entities

The last version (release 0.5) of the DRO is represented in RDFS, publicly available at [22], and is composed of 96 classes (9 inherited from CIDOC CRM and 5 from FRBRoo) and 44 properties (7 inherited from CIDOC CRM and 2 from FRBRoo).

An overview of the core entities is depicted in Fig. 2. A unique prefix is used to identify each concept. The prefix is composed of a letter that serves as a namespace identifier, and a number that identifies the concept. The letters are "E" and "C" for the concepts inherited from CIDOC CRM, "F" for those taken from FRBRoo and "D" for DRO specific concepts. The same for properties, "P" and "S" identify the ones inherited from CIDOC-CRM, "R" is used for those taken from FRBRoo and "I" for DRO specific properties. The final "F" means forward, while "B" means backward.

The main classes used to formalise the legal framework are *Regulation* and *NationalRightType*. Regulations describe patterns of situations that are permitted. This is what Rights Expression Languages aim to express and control. Regulation splits in two subclasses: *WrittenNorm* and *Agreement*. *WrittenNorm* represents all laws which are valid in a certain country at a certain time; *Agreement* describes both international agreements which override local laws and bilateral agreements between right holders and other people. A right holder may in fact transfer one or more rights that he owns through an *IPRContract* or grant some permissions to other people to act on his digital objects issuing them an *IPRLicense*. *NationalRightType* is used to represent the types of rights that have a validity in a given country at a given time.

As already pointed out, the DRO distinguishes between two kinds of rights: the ownership rights, which are the exclusive rights typically held by the authors of the creative works, and the permissions that are granted by the right holders to other persons to use such works. The class *OwnershipRight* models the first type of rights, while the class *Permission* includes all types of authorisations to make use of content, including the authorisations given by the law and those given by the right holders through licenses. In both cases it is a consequence of a *Regulation*, respectively a *WrittenNorm* and an *IPRLicense*.

The other core entities which are part of the DRO have been adopted from CIDOC CRM and FRBRoo, in particular the Creation and the Action models, i.e. concepts like *Actor*, *LegalObject* and *Activity* with their subclasses and relationships. These concepts have been linked to the concepts which are specific of the copyright domain through the use of suitable relationships. For instance an *Actor owns* a *Right* which *isOn* a *LegalObject* and a *Right allows* (or *disallows*) an *Activity*.

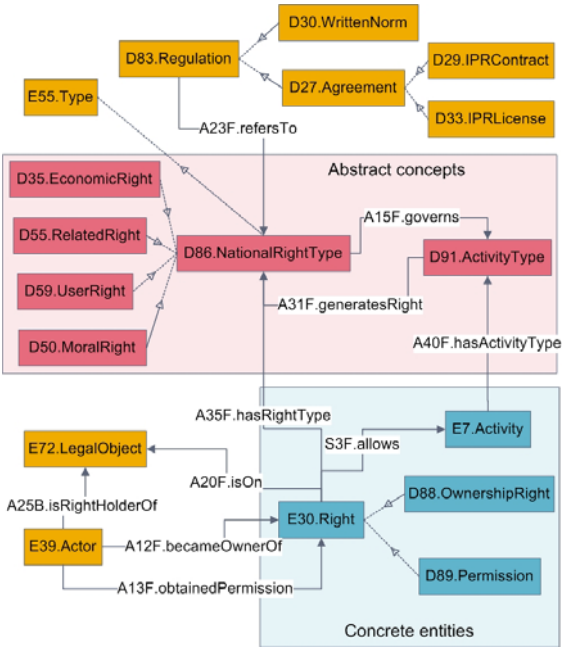

Fig. 2. Overview of DRO core entities

Another important characteristic inherited from the CIDOC CRM is the distinction between individual entities on one side, like persons, objects, licenses, etc., and general concepts on the other side, like types of rights, types of activities, constraints, and others, which are used to categorise individual entities. So, in the DRO there are more abstraction levels of rights entities: *Right* represents the instances of rights held by individual legal and physical persons on a precise object, while *NationalRightType* models types of rights. The same for *Activity* and *ActivityType*. This approach reflects looking also at the properties, so we have that a *RightType governs* an *ActivityType* and, at the corresponding individual level, a particular *Right allows* a specific *Activity*.

Therefore *OwnershipRight* and *Permission* represent rights of a given person on a precise object, while the pattern of situation that is allowed is represented by apposite entities, such as *ActivityType*, *PermissionPattern*, *Constraint*, *Condition*, *Validity* and others, together with properties such as *hasDuration*, *hasPurpose*, *hasExerciseLimit*, *hasAuthorisedPrincipals*, *hasFee*. Individual situations, expressed in terms of ontology instances, should then be matched against the general patterns, still expressed in terms of instances of the ontology.

Another important notion that was introduced is that of *Validity*, in order to address the problems due to the existence of different national legislations. Each *NationalRightType*, for instance the *DistributionRight* or the *AttributionRight*, is modelled as having a validity in space and time, and is linked to a set of applicable *Regulations*. Also the rights-generating activities, for instance *Derive*, *ProduceFixation* or *Perform*, have the dimensions of space and time as one of their attributes. This makes it possible to derive the precise instance of *NationalRightType* to which the individual ownership right corresponds, thus linking it to regulating Laws.

5 Examples

5.1 Identification and Reporting of Rights

The first application of the Digital Rights Ontology is to use it as a formal dictionary to express rights, including ownership rights and permissions.

Within the CASPAR Framework Architecture [23][24], a Digital Rights Management component has been developed that derives all the existing ownership rights given the provenance of a work. In particular, all relevant history information about a given work can be registered via a specific API, such as who performed which activity, thus contributing to the production of the overall work. On the basis of this provenance information and of the copyright law of a given country, the DRM component determines who holds any specific ownership right on any creative contribution to the work, for instance who holds the distribution right on the lyrics of a musical piece.

The DRM component can also export the information about the ownership rights that it has derived, and uses the DRO for that purpose. In the CASPAR project, this rights export service has been implemented to generate provenance metadata, namely a subset of PDI (Preservation Description Information), to be associated to the archived content data objects. It is of particular benefit for preservation purposes to have PDIs analytically expressed in a knowledge representation language, as this simplifies the preservation of the PDI itself. In fact, the ontology provides both a dictionary and a formal semantics of its entities.

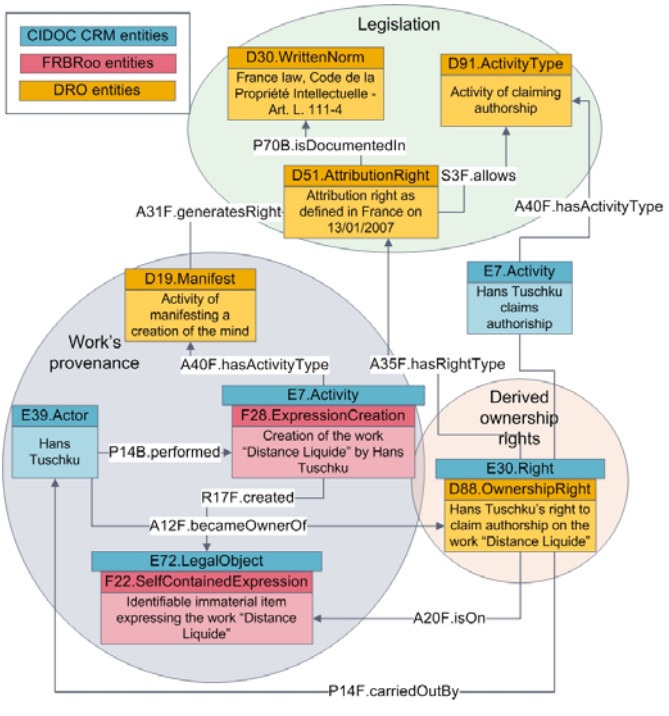

Fig. 3. Identification of an Attribution Right

Fig. 3 describes the pattern that is evaluated by the DRM component to derive a specific ownership right, namely the right to claim authorship on a given intellectual creation. The diagram highlights that there are two groups of entities that contribute in bringing an ownership right into existence. On one side there are the provenance events and related entities, which are specific to the work, and on the other side there are some more general applicable entities, defined by the legal framework of a given country. The latter express the rules by which certain types of activities generate rights-protected products, and the activity types on which the product creators have the exclusive right.

Looking at the figure, Hans Tutschku holds the exclusive right to claim authorship on the work "Distance Liquide", as a consequence that he composed this acousmatic work and that the copyright law assigns the attribution right to the creators of intellectual works.

5.2 Evaluation of Rights and Consequences of Actions

An innovative characteristic of the Digital Rights Ontology is that it makes the nature of rights explicit, namely that it expresses the permissions as patterns of constrained

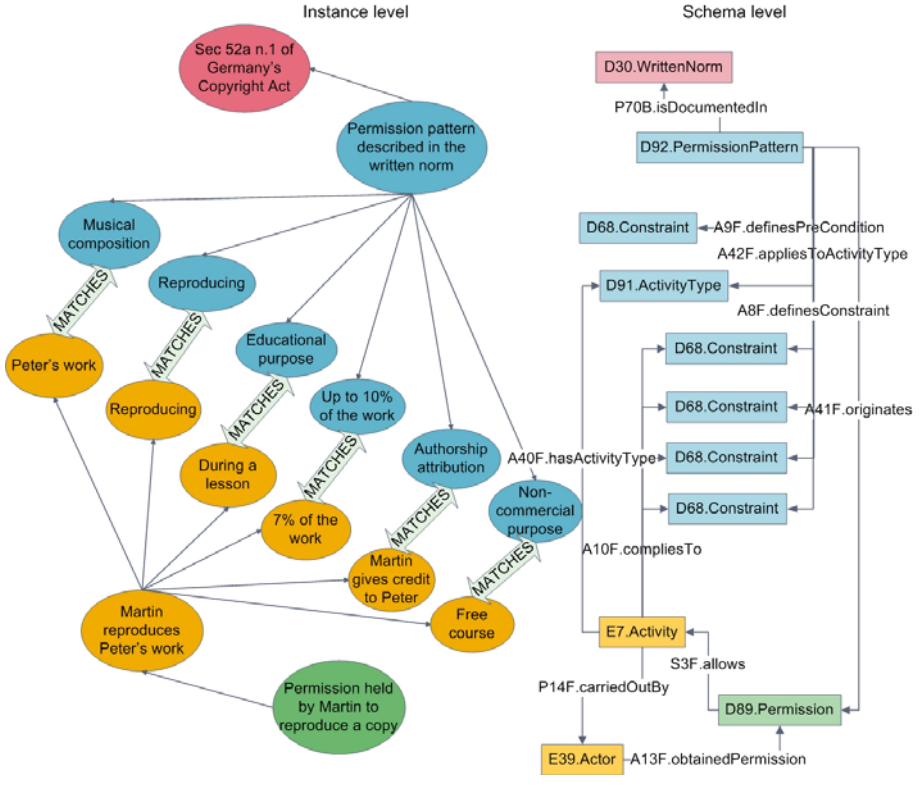

Fig. 4. Right evaluation as pattern matching

activities. Therefore the evaluation of permissions corresponds to the matching of real case situations against the patterns described in licenses or in written norms. This means that it is required from one side to be able to capture the details of the consumers' intended actions and on the other side to formulate the constituent parts of the permissions, be they given through licenses or by the law.

For instance, the educational user right can be modelled, likewise a license, as a set of *PermissionPatterns*, namely one for each type of activity. Fig. 4 depicts an example where the activity is the reproduction of a musical piece. The intended action is highlighted on the bottom half and the top half is the permission pattern. The individual right that the teacher Martin holds, allowing him to reproduce an excerpt of the musical piece during a lesson, is derived from the successful matching between the instances describing the intended action and the instances describing the pattern. Finally, this *PermissionPattern* is bound to a set of *WrittenNorms* that regulate it, which also declare the consequences of violating it.

Yet, in many cases it is difficult to implement the automatic execution of such a pattern matching process. The idea is that certain conditions would be cut out from the evaluation process, and would just be presented to the consumer. Furthermore, the norms related to the particular use case could be referenced.

So far, only the concepts necessary as parameters for this rights evaluation process have been identified. A related implementation is foreseen by procedural code.

6 Conclusions

The main purpose of the Digital Rights Ontology is to provide a standard model that allows digital repositories to identify existing rights on digital objects and to evaluate the consequences of potential actions which may fall under IPR-related regulations. This is achieved by making explicit all the fundamental aspects and events in the history of a work that influence the status of rights (i.e. digital provenance, licenses and legal framework) and by interpreting the rights as patterns of future activities.

This innovative idea also represents a key aspect for the purpose of determining, preserving and enforcing digital rights in the long term. In this perspective the DRO is being used, within the CASPAR Framework, as the reference conceptual model for digital rights. The implementation of the Digital Rights Management component relies upon it as the formal language for exporting information about ownership rights, once they are derived by applying the copyright norms of a particular country to the digital provenance information of a given work. The logics of the respective evaluation has been implemented in procedural code operating on a database, which addresses the scalability requirements of the CASPAR system better than a rule-based system, but this is not the subject of this paper. The ongoing status of development can be followed on [24].

The compatibility of the DRO with important ontologies for the description of intellectual work - CIDOC CRM, FRBRoo - makes it a promising candidate to support interoperability of potential global digital rights clearance services across on-line digital repositories.

Future activities related to the DRO include the prototype implementation of a rights enforcing algorithm, so far foreseen by procedural code, that would be based on

the pattern matching model described in this paper. In principle, the pattern matching could also be done by KR reasoning systems, but for reasons of platform restrictions and performance concerns, we have not yet considered this approach.

In addition, the digital provenance sub-model and the derivation of rights will be further validated, in particular with respect to the challenging cases of incorporation and derivation of works, where determining the rights is particularly complex.

Acknowledgments. We would like to express special thanks to the European Commission for funding our work in the fields of Digital Rights and Ontology and to the useful cooperation of the research team of the CASPAR project, in particular David Giaretta, project coordinator. Thanks to Fiore Basile who conceived the work and to David Lamb and Francesca Fenili for their feedback for improving the presentation.

References

1. Delgado, J., Gallego, I., Llorente, S., García, R.: IPROnto: An Ontology for Digital Rights Management. In: Frontiers in Artificial Intelligence and Applications. 16th Annual Conference on Legal Knowledge and Information Systems. JURIX 2003, vol. 106, pp. 111–120. IOS Press, Amsterdam (2003), http://www.jurix.nl/pdf/j03-12.pdf
2. Nadah, N., Dulong de Rosnay, M., Bachimont, B.: Licensing digital content with a generic ontology: escaping from the jungle of rights expression languages. In: International Association for Artificial Intelligence and Law, pp. 65–69 (June 2007), http://doi.acm.org/10.1145/1276318.1276330
3. Coyle, K.: Rights expression languages - a report for the library of congress (February 2004), http://www.loc.gov/standards/relreport.pdf
4. Doerr M., Bekiari C.: FRBRoo, a Conceptual Model for Performing Arts. In: 2008 Annual Conference of CIDOC, Athens (September 2008), http://www.cidoc2008.gr/cidoc/Documents/papers/drfile.2008-06-18.1838425657
5. Multimedia Description Schemes (MDS) Group: Introducing MPEG-21 REL - an Overview. International Organization for Standardization (July 2005), http://www.chiariglione.org/MPEG/technologies/mp21-rel/index.htm
6. Iannella, R.: Open Digital Rights Language (ODRL) Specification. Version 1.1 (August 2002), http://odrl.net/1.1/ODRL-11.pdf
7. Logos and codes for the CC core licenses, http://creativecommons.org/about/licenses/meet-the-licenses
8. Description of ccREL, http://wiki.creativecommons.org/images/d/d6/Ccrel-1.0.pdf
9. Overview of ccREL vocabulary (RDF Schema), http://creativecommons.org/schema.rdf
10. Definition of Copyleft licensing, http://en.wikipedia.org/wiki/Copyleft
11. García, R.: A Semantic Web Approach to Digital Rights Management. PhD Thesis, Universitat Pompeu Fabra, Barcelona (2005), http://rhizomik.net/~roberto/thesis/
12. García, R., Gil, R., Delgado, J.: A web ontologies framework for digital rights management. In: Artificial Intelligence and Law, vol. 15(2), pp. 137–154. Kluwer Academic Publishers, Dordrecht (2007)

13. Bide, M., Rust, G.: The <indecs> metadata framework: Principles, model and data dictionary. Godfrey Rust, MUZE Inc., Mark Bide, EDItEUR (June 2000), http://www.doi.org/topics/indecs/indecs_framework_2000.pdf
14. CIDOC CRM v0.4.2, http://cidoc.ics.forth.gr/official_release_cidoc.html
15. Doerr, M.: The CIDOC CRM – An Ontological Approach to Semantic Interoperability of Metadata. AI Magazine 24(3) (2003), http://cidoc.ics.forth.gr/docs/ontological_approach.pdf
16. CIDOC CRM Web Site, http://www.cidoc-crm.org
17. Doerr, M., Le Bœuf, P., Bekiari, C. (eds.): International Working Group on FRBR and CIDOC CRM Harmonisation: FRBR object-oriented definition and mapping to FRBR ER (version 0.9 draft) (January 2008), http://cidoc.ics.forth.gr/docs/frbr_oo/frbr_docs/FRBR_oo_V0.9.pdf
18. FRBRoo v1.0, http://cidoc.ics.forth.gr/frbr_drafts.html
19. Le Bœuf, P.: Functional Requirements for Bibliographic Records (FRBR): Hype or Cure-All? Haworth Information Press Inc. (2005) ISBN: 0789027984
20. Doerr, M., Hunter, J., Lagoze, C.: Towards a Core Ontology for Information Integration. Journal of Digital Information 4(1) Article No. 169 (April 2003)
21. Smith, B.: Ontology: An Introduction. In: Floridi, L. (ed.) The Blackwell Guide to the Philosophy of Computing and Information, pp. 155–166. Blackwell, Oxford (2003)
22. Digital Rights Ontology v0.5, http://www.casparpreserves.eu/publications/ontologies/RightsOntology
23. CASPAR project Web Site, http://www.casparpreserves.eu
24. CASPAR software, documentation and code, http://developers.casparpreserves.eu:8080/