

Cognitive Interfaces for Legal Expressions Description – Application to Copyrighted Works Online Sharing and Transactions

Melanie Dulong de Rosnay

*Research Centre for Administrative Science (CERSA)
Information Technology, Law and Linguistic Laboratory
CNRS - Law University Paris II, Paris, France
melanie.dulong-de-rosnay@medialive.fr*

Abstract. Legal and technical means regulate the digital dissemination of creative works. Technological protection measures and rights management information implementing copyright licenses or contracts are based on Rights Expression Languages, i.e. machine-readable ontologies that link copyright concepts and terms to events occurring on digital networks. While developing such devices, specific attention shall be dedicated users' expectations in terms of problem description and legal situations expression. In order to be interoperable and fair, Digital Rights Management systems need a common vocabulary describing legal use cases and supporting opposing interests, as well as architectural conditions enabling to describe contexts and express questions, toward a dialog between users. After a presentation of the domain legal framework and an evaluation of current standardization initiatives, we will apply the concept of cognitive interface to the conception of copyright regulation applications.

1 Introduction

The combination of compression, network capabilities and digital rights management technologies offers a great opportunity for a wide, quick and low cost public dissemination of creative and information works. Copyrighted works management is getting highly automated and IPMP systems are becoming Information Technologies regulators. Rights holders expect DRM systems to express and enforce legal conditions under which they decide to disseminate their works. Intermediate and end-users with no legal education expect to access to works through user-friendly and interoperable interfaces.

These interfaces are intended to describe contractual and legal situations. Both the vocabulary and the structure of these interfaces shall be compatible with accurate legal expression within copyright legislation requirements, enabling use context description by keywords or metadata simple to understand for all users, from legal professionals to non legal-experts.

Cognitive interfaces are defined as modes of questioning based on dynamic shape of users requests rather than on problem domain static knowledge. Decision making models including IPMP architectures and DRM applications can be build upon a representation of legal concepts and procedures. We therefore propose to develop a cognitive interface, which will allow the communication between various software resources and agents, "taking into account a meta-knowledge implied in the situation of dialog with a view to adapting it to types of specific requests in a domain" [2].

2 Users Expectations about Digital Copyright Management Systems

The term ‘user’ is understood as all groups of actors of the dissemination chain interacting with copyrighted works: authors, creators, producers, distributors, intermediate and end-users (private, corporate and institutional), broadcasters...

Authors, original creators of works distributed on networks, have the exclusive right to authorize the reproduction and the communication of their works, within certain limits. They might license these rights to producers or mandate collective societies for the commercial exploitation of their works.

Authors benefit from technological innovations at several levels. They will find a large audience for their works in an environment where several barriers are removed. Audiences are no longer restricted to national markets and linguistic areas; authors can benefit from distribution platforms, providing a large visibility, metadata for an enriched access to content according to end-user’s preferences; authors can also publish and distribute directly their works through personal websites or community portals.

They expect IPMP systems to express and enforce the rights they decided to attach to their works. Their expectations can be financial but also related to the moral control of works’ destination, initiatives close to open source philosophy being “devoted to expanding the use of creative work available for others to build upon and share” [25]. Identification, marking and watermarking systems ensure a persistent protection of their moral rights, at first attribution rights i.e. paternity recognition. Authors also expect to enjoy an easy access to other authors’ works as source of inspiration, or as reusable material to be integrated to their own works, according to new creation techniques (sampling, VJ, cut-up, tributes, covers...). Information Technologies can provide interfaces facilitating contacts between creative communities; and DRM are expected, if not to support, at least not to restrict those exchanges.

End-users expect a user-friendly and easy access to culture, education, heritage and entertainment while preserving their privacy. Producers and distributors are looking for a return on investment, reliable event reporting, and cost-effective management. Producer’s representatives are afraid of peer-to-peer file sharing consequences on their sales. They express publicly their disapproval by asking for Internet service technical providers and Courts support. Public institutions, open source communities and academics fear that DRM systems, acting as electronic fences, replace legal public mediation and restrict exceptions to exclusive rights and therefore legitimate access to works.

3 Technical Standardisation and Metadata Interoperability

Electronic commerce is faced to obstacles that prevent the full development of digital media and services. Digital media creation, sharing and distribution has to be supported by intelligent DRM systems, able to adapt to new use models and to cope with contradictory interests of all stakeholders. DRM solutions should therefore be able to process dialog, arbitration, mediation between those expectations expressions. It should also allow easy update of content rights management policy, according to expectations and time-to-market strategies evolution.

From Consumer Electronics manufacturers and industry point of view, bits shall be delivered at low cost, technical standardization of delivery formats enabling interoperability between systems and devices. The Moving Picture Expert Group [17] develops ISO technical standards for coded representation of digital audio and video, and enabled the development of multimedia standards such as mp3 format, digital television, DVD, multimedia information retrieval and description... Launched in June 2000, MPEG-21 standard [18] plans to define a global framework to enable transparent and augmented use of multimedia resources

across networks, equipments and communities through an architecture structuring identification, management and protection of multimedia works. Software tools shall provide interoperability enabling to use the content “without being bugged by incompatible formats, codecs, metadata” [19].

The framework foresees therefore to include the definition of legal metadata identifying works and describing rights, users and conditions. MPEG-21 is expected to make interaction with content more transparent, and to define the relationships between the normative technologies needed to regulate the multimedia delivery chain. These elements are elaborated by defining the syntax and semantics of their characteristics, such as interfaces to the elements.

The integration of those modules is not only faced to technical requirements, but also to emerging uses and business models. A machine-readable language able to express rights, which are getting more and more complex in the digital world, is supposed to define algorithms enforcing electronic content governance within a trusted environment. “Code is Law”, and software developers shall be aware of the values they embed inside their code and architectural design [9], [10]. Computed access control may indeed stand for legal code. Current DRM systems are based on interoperable languages describing terms and conditions for digital resources manipulation, transport and usage. Such languages shall be able to express simple and complex rights expressions, based on standardized Rights Data Dictionaries. An IPMP language and dictionary will be used on IPMP interface to describe use context, users, actions, rights, works, conditions...

4 Rights Expression Languages State of the Art

Grounded on Imprimatur European project results [21], <indec> and <indec>2rdd [20] provide a methodology for the interoperability of metadata’s supporting electronic commerce. Rights and permissions basic structures are expressed by a Context Model defining the relationships between entities, organized around the notion of Action on a Resource.

ODRL (Open Digital Rights Language) [22] developed another semantic on XML standard. Rights governing Assets are including Permissions characterized by Constraints, Requirements and Conditions. Rights are expressed by Parties through Offers, Agreements and Revoking Rights.

XrML, Extensible Rights Markup Language [23] is organized around Authorization Context fundamental notion: a Right can be performed by a Principal on a Resource under specific Conditions.

IPRonto [24] formalized on a semantic web approach, structures IPR information at a higher level of abstraction. This ontological domain representation expresses IPR roles and events capturing the various steps of the content lifecycle: creation, rights transfer, distribution, use, reproduction, and transformation, providing a broad framework for Automatic Negotiation. Two consenters are involved, the Requester and the Licensing Agent, representing the rights holders. Offers and Counteroffers expressed with IPRonto help to reach an Agreement for a defined time period, compensation and action [4].

5 Legal Requirements for IPMP Architectures

Our position elaborated in context of MPEG-21 requirements [19] claims that in order to be agnostic and not to preclude non-static legislations and favour specific business models, only a minimum core of compulsory features (the lowest common denominator of worldwide copyright legislation and contractual practices) should be specified within a standard that aims at defining an interoperable framework for future DRM applications.

Legal knowledge has to be embedded in an interface expressing copyright law context and questions related to content distribution: licensing, infringement, and exception.

We will first present legal requirements with regards to the hierarchical relationship between international conventions and national legislations. Legal norms provide their own exceptions: prior rights holder's authorization is requested before use, but a work unauthorized use do not always constitute a copyright infringement. We will then describe the consequences of those legal requirements on IPMP architecture, and propose several architectural features. The exclusive right of authorizing can be materialized within an automated dialog through an 'Event Request' made by the client to the rights holder or its representative. Communication between users is made possible through the implementation of a Rights Expression 'bi-directional dialog', formalizing contractual conditions.

5.1 *National Copyright Expressions*

National copyright expressions allocate different rights to entities, and these entities are defined differently. For instance, works categorization has implications on work's legal qualification (audiovisual or multimedia work, database or software, are handled differently) and similar works will receive different treatments and induce different remunerations according to the jurisdiction they depend. These local differences have consequences on rights expressions representation. Negotiation conditions will vary from one case to another and jurisdiction is not always simple to determine in case of international sales and international law conflicts. Reasoning automated systems (e.g. authorization, compulsory licences and royalties sharing, fair use scope) shall be adapted, as embedded concepts may have different interpretation and deserve different processing.

5.2 *International Legal Principles: the Authorization/Exception Pattern*

An analysis of legislative texts and situations allowed the retrieval of copyright situations patterns. The model of prior authorization for limited rights granting is associated to the principle of exception to exclusive rights, applying in certain cases determined differently in each jurisdiction according to the use context

Some principles or pieces of legislation are internationally recognized. At the top of copyright norms international hierarchy, WIPO Berne Convention [13] is enforceable in most countries. Its minimum common legal principles are translated in national copyright laws and shall therefore be embedded into copyright management systems. 'Exclusive rights' means that rights holders are, if no other provision is applying, sole actors allowed to fulfil or authorize the accomplishment of specific acts related to their works. Even if it is not always desirable [11], prior authorization from the rights holder shall usually be requested before any analogue or digital action, such as digital display, reproduction, communication to the public, performance, distribution, rental, translation, adaptation... Authorization is then granted for a defined use context in exchange of remuneration.

Nevertheless, this general principle presents several exceptions within international and national legislations that avoid the user to have to ask rights holders' prior authorization every time before s/he wishes to perform an action on a protected content. Before XVIIIth century legislations limiting uses by a temporary monopoly, creative works natural state was full availability. Public domain was the general rule, not the exception. Current prior authorization legal paradigm could be reversed: some use might be freed and exclusive or reserved rights would become the exception.

Limitations on exclusive rights have various explanations (technique, privacy, fair use...) and restrict the obligation to request an authorization before performing an action. Temporary acts of reproduction part of a network transmission technological process are exempted from prior authorization in Europe, but not in USA. Favourable translation licensing conditions can be granted to developing countries [13]. Compulsory licenses avoid asking prior authorization before any act, because it is burdensome, materially impossible (radio broadcasting...), or mandatory (cable distribution in Europe) to proceed through collective management. Authors can also restrict their exclusive rights on a voluntary basis and place their creations into the public domain, or voluntarily restrict their exclusive rights to commercial uses [25]. Fair uses are generally admitted if they do not affect normal exploitation or rights holders' legitimate interests, but it is up to national states to precise their scope and implementation.

If an exception to exclusive rights applies, an action that would deserve prior authorization is not considered as an infringement and rights holders shall on no account restrict content access or usage through DRM. But an exception can be enforced, meaning an event will not require prior authorization, only once the use case context has been defined and qualified, and not previously according to content nature or user affiliation. At a first glance, it looks impossible to formally represent a priori fair use cases. Nevertheless, it is important to distinguish legal authorization concept from client-server automated processing authorization requests. Legal concepts or rights and their implementation in a computer system belong to different knowledge levels and do not necessarily match point-to-point.

6 A Cognitive Interface for Problems Expression and Solving

6.1 Expression of Events, Requests and Computed Actions Based on Rights Granting Dialog

Notwithstanding limitations on exclusive rights prerogatives, we can assume that all DRM applications and REL systems will translate prior authorization requirement for commercial licensing. After an authorization request, rights holders will answer by a minimal Boolean operation: refuse or allow the requested act under previously stated conditions. The agreement is typically materialized through a temporary licence agreement sent back from the rights holder to the client and activating rights processing on the end-user terminal, i.e. reproduction right will be materialized by the operation 'print' or 'copy'. Of course, there are valid models and practices that work differently from this general way, and architecture requirements shall, rather than limit choices, try to satisfy the broadest scope of requirements. If proposed licensing terms do not fulfil user expectations, s/he should have the opportunity and the tools to express her/his own particular conditions and confront them online to the rights holder disputed licensing terms.

6.2 Adaptation to Use Profiles through Scalability and Subsidiarity

Most client-server architectures allow a dialog between users and the added value of a cognitive interface is to provide the terms to express both rights and processing of actions on digital works files. Not only rights holders shall be allowed to express and propose use conditions, leaving end-users with no other choice than approving all the licensing terms or refusing them all and ending the transaction. The end-user shall be made able to propose alternative Rights Expressions and confront them to rights holder's Rights Expressions.

A cognitive interface will not propose the whole set of terms and rules, burdening the system and the user, but only the knowledge needed for the problem. The knowledge database

shall therefore be scalable and propose different sets according to use cases or users profiles. The situation can be described with many criteria. The categorization of the action (referring to the Right) and of the place (human-readable definition for applicable law) can be expressed easily in simple cases. On the contrary, the categorization of the user him/herself may request the interaction of a trusted third-party to validate the user legal qualification, i.e. if identity and location have legal consequences.

The range of implicated criteria determining the set of expressions and terms proposed to the user is dynamic, such as the level of control and trust. Sometimes, decisions shall be taken at a high level whereas in other cases, they can or even need to be taken at a lower level. Subsidiarity concept might help to determine the ad hoc level where a policy will receive the best management, and to distribute competences between decisional architecture levels, deduced from an efficiency assessment. Subsidiarity principle was first expressed in XIIIth century by Thomas Aquinas and recalled by constitutional federalists (Germany, Swiss, Italy, European Community...). High-level entities' action is legitimated only and as long as their added value enhances the inadequate capacity of lesser entities. Subsidiarity can be vertical and regulate governance levels. It can also have horizontal applications and help to share competences between regulating actors (public/private, law, standard, market, individuals, and even elements of an abstract architecture). As far as and DRM architectural and interface design are concerned, this concept might also help to choose the right level of management and decision.

6.3 *An Architecture for Rights Expressions Local Adaptation*

The mass of Rights Expressions combinations makes it impossible, regarding networks and users congestion, to associate all possible legal national requirements to a single request or piece of content. Therefore, it is advisable to delegate Rights Expression local governance at a lesser level than worldwide content provider. We consider here the possibility for a central content server to delegate legal questions to specialized trusted servers. A local server contacted by a central server will process Rights Expression Adaptation to non-static local legal requirements.

A first central server would send the resource and its Rights Expression to many other decentralized servers, which would adapt it to local requirements. Such a second level server can be a specialized network or a national platform. If it is dedicated to private documents sharing within a family, it will offer the broadest scope of Rights Expressions: all actions being allowed, and no authorization or reporting mechanisms between registered members preventing access to content. On the contrary, if the second level server is a commercial application, it will request a fee for a determined event, all other events being forbidden. Providing a scalable range of applicable terms and licenses would permit to handle every situation with the adequate level of control. Any user and right holder would choose the set corresponding to its needs and expressed situation. Each level would define its own rules and desired level of control on actions according to legal, business, social norms and uses.

A first Rights Expression is requesting a prior authorization (1), to be applied if nothing else indicates a fair use or public domain context. Local servers adapt (4), the first Rights Expression to national laws e.g. mandatory levy or classification update and local business model. Prior authorization condition will be skipped if the Rights Expressions from the top meets a Rights Expressions from a lower level that exempts the specific action from prior authorization (public domain, compulsory licence, fair use national provision...).

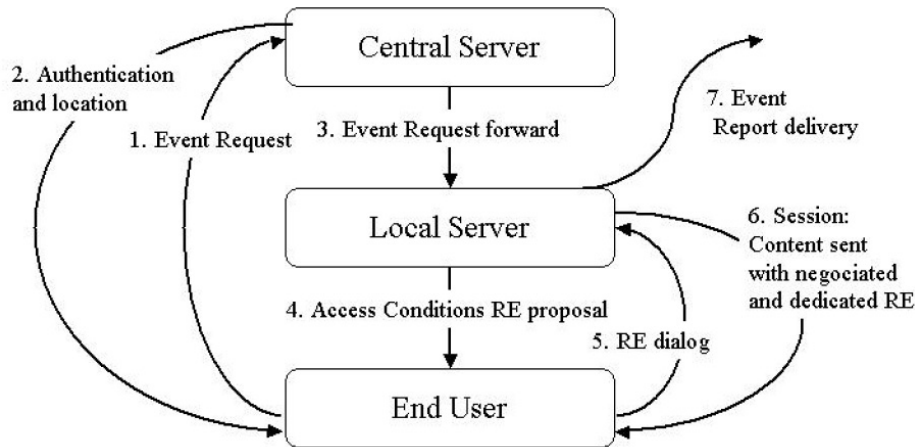


Figure 1: Proposed IPMP architecture enabling Rights Expressions (RE) dialog.

7 Current and Future Work: a Knowledge Database for Copyright Online Problems Description

A knowledge database gathering possible terms and concepts for Rights Expressions is being elaborated. According to intended use, and in order to allow a dialog between users and levels, Rights Expressions have to be formally represented by an ontology to express questions relevant to the context and propose answers corresponding to users' needs. Rights Expressions have to support confrontation, negotiation, decision and solution enforcement (licence agreement, copyright exception, nothing) between Rights Expressions set up by rights holders and use conditions adapted by a local server. The benefit of including a dialog allows real world dialog conditions between users. Communication is allowed in both directions (i.e. with a return path) and not only reduced to an offer from the rights holder representative to the end-user with no further choice than block acceptance or refusal. If the Rights Expressions issued by the right holder do not foresee to grant printing permission, the end-user might well communicate his request to the server.

7.1 A Methodology to Build a Legal Ontology

Even if efforts have been made by 1996 international treaties [15], [16] and 2001 European Directive [14], copyright national terms are not harmonized. Their interpretation is subject to changes and uncertainties, which are unfavourable to investment and commerce development. Legal concepts are open textured concepts [8] as their understanding is related to external conditions such as the context. This diversity generates difficulties for the implementation of DRM systems. Indeed, it is important to express the conditions attached to a work's use with accuracy and flexibility.

Ontologies can be defined as referential abstractions providing notions and concepts of a domain with a formalization of the links between those concepts [7]. This modelling is useful for knowledge representation and computed exploitation. Our work-in-progress research is based on the development of a legal ontology that will allow expressing rights and use situations attached to a context and a work. As it is possible to build several ontologies for a single legal domain, the composition of our ontology will be deduced from the task it is designed for [1], i.e. expression of questions and decisions about copyright and use context description by all users, balancing stakeholders interests within creative content electronic delivery. The cognitive interface and its knowledge database should be reusable by other IPMP and DRM

systems.

The preliminary phase of our work was to study previous results of other Rights Expression Languages (RELs) and dictionaries, most relevant being XrML, ODRL, <indecs> and IPRonto. Except the latter one, all projects started from use cases and not from legal texts in order to define a core set of semantics.

The added value of the proposed methodology is to start from an expertise of the legal knowledge domain, defined as the intersection between copyright law, law for Information Technology and media law. This expertise aims at avoiding texts partial understanding or misinterpretations, as legal doctrine and case law knowledge enable to throw light on texts interpretation. Our reasoning process aims at modelling the core of all use cases and especially re-use cases, the latter aspect being outside other projects' scope. We also include public institutions requirements instead of limiting our representation to business interests. Archival or re-use of protected content in a library, a museum or within teaching and research materials are definitively part of content electronic delivery targets and shall be taken into account [6], [12]. Other REL give priority to commercial uses and hardly tackle fair use because it is technically easier to close than to open access. Making a work available for specific uses may cause security wholes for unauthorized persons. REL are deemed to be agnostic and able to express all uses and legislations, at least indirectly or by default. However, a statistical study of the vocabulary captured by existing terminologies allowed to detect an important under-representation of terms and concepts attached to end-users and copyright exceptions, in comparison with expressions describing rights holders, collective management, contractual conditions and remuneration sharing.

We propose to restore copyright initial purpose and are developing a bank of terms large and representative enough to describe core situations and specificities with respect to un-governed uses. A cognitive interface should process access requests according to context and express use cases with a simple metadata system, reducing the situation to the ad hoc level of legal information needed for the problem description and solving. Applications for a cognitive interface providing appropriate semantic (rights expressions) and syntax (computed architecture for questions fields) can be selection of relevant text to allow or turn down a request, rights holder license granting and remuneration process automation. The metadata system, based on a copyright ontology, is supposed to enable non-lawyer users to describe their status and requested use conditions.

7.2 First Classes and Concepts

An ontology dedicated to copyright situations description and expression shall be able to describe at a meta level national copyright legislations and case law subtleties. Classes shall be intuitive enough for a self-categorization by non-lawyer users. Links between classes (users, rights, attributes, decisions...) are representing digital content life-cycle steps from creation to ungoverned and non-copyrightable uses (public domain, fair use...) to licensing and re-use management.

Here is an excerpt presenting terms categorization. Terms have been extracted from legal texts and existing Rights Expression Languages with the objective of reflecting use cases and user profiles. Subsequent intensive work is dedicated to the completion of the domain ontology, including concepts definition in national legal systems and links design between **classes**, concepts and associated *attributes* in order to implement it within a DRM system.

Domain conceptualisation within an ontology can help users to express a situation and an action, but also to appreciate an existing DRM system and evaluate its completeness and embedded values, control, trust or privacy level. This aspect can be helpful when a user has

Table 1: Excerpt of terms classification prior to ontological domain representation

USER	WORK	RIGHT-ACT	CONDITION	DECISION
Licensor	Non copy-righted	Economic Rights	<i>Fee</i>	Fair use case identification
Right Holder	Public domain	Communicate	<i>Quantity</i>	<i>Berne Convention</i>
Author	Ideas	Copy	<i>Levy</i>	<i>three-step test</i>
Producer	Personal data	Distribute	<i>Time</i>	<i>Quotation scope criteria</i>
Distributor	Copyrighted	Derivate	<i>Format</i>	License agreement formalisation
Collective society	Book	Aggregate	<i>Quality</i>	<i>Resolution</i>
Licensee	Translation	Quote	<i>Location</i>	<i>Public</i>
End-user	Joint work	Rent	Moral rights	<i>Exploitation domain</i>
Private user	Phonogram	Lend	Attribution	Copyright infringement
Teacher	Motion picture	Moral rights	Integrity/Respect	Access refusal
Journalist	Collective work	Withdrawal	Open source	
	Videogame	Divulgate	Copyleft	
	Software			
	Database			

to choose between different applications reflecting an IPMP policy.

Acknowledgments

We would like to acknowledge CERSA and MEDIALIVE for supervising and funding our research within MPEG-21 stimulating framework.

References

- [1] Trevor Bench-Capon, Task Neutral Ontologies, Common Sense Ontologies and Legal Information Systems. Second International Workshop on Legal Ontologies, JURIX 2001, Amsterdam.
- [2] Filipe Borges, Danièle Bourcier, Evelyne Andreewsky, Raoul Borges, Conception of cognitive Interfaces for legal knowledge – Evolution of the JURISQUE project on the risks of avalanches, ICAIL 2001 The Eighth International Conference on Artificial Intelligence and Law, May 21-25, 2001 St. Louis.
- [3] Jaime Delgado, Isabel Gallero, Standardisation of the management of Intellectual Property Rights in multimedia content, Proceedings of the Second International Conference on Web Delivery of Music, IEEE Computer Society, 2002, p. 125-132.
- [4] Séverine Dussolier, Fair use by design in the European Directive of 2001: an empty promise, 12th Conference on Computers, Freedom and Privacy, 2002. <http://www.cfp2002.org/fairuse/dusollier.pdf>
- [5] Tom Gruber, A Translation Approach to Portable Ontology Specification, Knowledge Acquisition, 5/1993, p. 199-220.
- [6] Guiraudé Lame, Using text analysis techniques to identify legal ontologies components, ICAIL 2003, Workshop on Legal Ontologies & Web Based Legal Information Management.
- [7] Lawrence Lessig, Code and Other Laws of Cyberspace, Basic Books, 1999.
- [8] Lawrence Lessig, The future of Ideas, The Fate of Commons in a Connected World, Random House, 2001.
- [9] Jessica Litman, Revising Copyright Law for the Information Age, 75 Oregon Law Review 19, 1996. <http://www.law.cornell.edu/commentary/intelpro/litrvtxt.htm>
- [10] Deirdre Mulligan, Aaron Burstein, Implementing Copyright Limitations in Rights Expression Languages, 2002 ACM Workshop on Digital Rights Management.

Legal texts

- [11] Berne Convention for the Protection of Literary and Artistic Works <http://www.wipo.int/clea/docs/en/wo/wo001fr.htm>
 - [12] Directive, 2001. Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, OJ L 167 22.06.2001 p. 10. <http://www.europa.eu.int/eur-lex/en/index.html>
 - [13] WIPO Copyright Treaty, World Intellectual Property Organization, 20/12/1996 <http://www.wipo.org/eng/diplconf/distrib/94dc.htm>
 - [14] WIPO Performances and Phonograms Treaty, World Intellectual Property Organization, 20/12/1996. <http://www.wipo.org/eng/diplconf/distrib/95dc.htm>
- DRM standardisation and REL/RDD initiatives**
- [15] MPEG, <http://www.mpeg.telecomitalia.com>
 - [16] MPEG-21 <http://www.chiariglione.org/mpeg/index.htm>
 - [17] MPEG-21 N4518. ISO/IEC JTC 1/SC 29/WG 11 N4518, From MPEG-1 to MPEG-21: creating an interoperable multimedia infrastructure <http://mpeg.nist.gov/>
 - [18] Indecs <http://www.indecs.org/pdf/framework.pdf>
 - [19] Imprimatur project <http://www.imprimatur.net>
 - [20] ODRL <http://odrl.net/>
 - [21] XrML <http://www.xrml.org/>
 - [22] IPROnto <http://dmag.upf.es/ontologies/ipronto/index.html>
 - [23] Creative Commons <http://www.creativecommons.org/>