From Narrative Contracts to Electronic Licenses: A Guided Translation Process for the Case of Audiovisual Content Management

Víctor Rodríguez, Jaime Delgado, Eva Rodríguez
DMAG (Distributed Multimedia Applications Group)
Universitat Pompeu Fabra, Passeig de Circumval•lació, 8, 08003 Barcelona, Spain
Universitat Politècnica de Catalunya, Jordi Girona, 1-3, 08034 Barcelona, Spain
victor.rodriguezd@upf.edu, jaime.delgado@ac.upc.edu, evar@ac.upc.edu

Abstract

Electronic licenses can represent the relevant information conveyed in narrative contracts. A standard Rights Expression Language (REL) license is extended here to represent the typical clauses found in the audiovisual market contracts, and a method is proposed to make this conversion. The extension is needed since the current REL does not support all the required rights and conditions. The translation process is done in two steps; in a first level the contract is organized in a structured XML format, and in the second level an electronic license is produced. This is done with a computer guided process whose resulting electronic license will be able to govern in the required manner audiovisual content distribution systems.

1. Introduction

Contracts are the expression of business agreements. In the context of e-commerce much has been studied about electronic versions of narrative contracts, but a standard digital format has not yet been established.

In the commerce of digital goods with Intellectual Property (IP), REL licenses have been used with success in content distribution platforms. A license expressed in languages like MPEG-21 REL [1] or ODRL [2] is able to represent precisely the information to govern and manage the digital rights but, nevertheless, its completeness is not guaranteed for our purpose. This paper presents a proposal, in the context of audiovisual market contracts, for *digital licenses* to include more information that is usually present in the traditional paper *narrative contracts*.

It has to be noted the accepted sense of the words *contract* and *license*. While a license is a permission to do an act that, without the permission, would be unlawful, a contract law, by contrast, revolves around the notion that two (or more) parties have bargained or negotiated an exchange of promises. This later notion includes the concept that one of the contracting parties has made an "offer" and the other party has accepted the offer, possible making a counterproposal. In this sense, throughout this paper we should be exclusively using the word “contract”, but we have respected also the term “license”, typical of REL vocabulary even when the agreement has been reached after negotiation.

2. Audiovisual Electronic Contracts

2.1 State of the Art on Electronic Contracts

The literature on digital contracts is extensive, defining its format, its lifecycle and its negotiation [3][4][5].

A good system for negotiating and executing contracts, COSMOS [6], was already introduced in the nineties; the contract was modeled and described in UML, and a CORBA distributed system led it into practice. This was presented before the XML era, and a new format would be most probably under the form of a XML Schema or a DTD. In fact, the alternatives that came later followed this idea, even progressing from the syntactic representation level to the semantic one [7][8][9]. The Content Reference Forum developed the Contract Expression Language (CEL) [10] as a XrML [11] based language that enables machine-readable representation of typical terms found in content distribution contracts, being compliant with the Buisness Collaboration Framework (BCF) [12].

Later on, and of capital importance, OASIS [13] established the LegalXML eContracts Technical Committee to evaluate a possible eContracts Schema, approving a first version of the standard on April 2007. [14].

But, on the contrary of these languages with ambitions of general use (excepting CEL), what we

0-7695-3030-3/07 $25.00 © 2007 IEEE
DOI 10.1109/AXMEDIS.2007.18

present only considers representing the B2B contracts in the audiovisual market. This will be achieved through an extension of the MPEG-21 REL license format, but before proceeding with the details, we will analyze the requirements of electronic contracts for our case.

2.2. Audiovisual Contracts Analysis

Contracts are made of a set of clauses plus certain metadata (signature, date of the agreement, signatory parties).

In the commerce of audiovisual products, parties are usually two: one party provides the audiovisual resource (generally called *licensor*) and other party makes use of it (the *licensee*). This use can be distribution, adaptation, any kind of exploitation or the mere enjoy of the work (this later case being the B2C case, of minor interest for our purposes).

Clauses are one or more paragraphs expressing unitary ideas, which according to the typical deontic logic, can be classified in rights, obligations and prohibitions, as can be seen in the next table:

	MAY	***MUST and MUST NOT***
Licensor	Rights of the licensor	Obligations and prohibitions of the licensor
Licensee	Rights of the licensee	Obligations and prohibitions of the licensee

Table 1. Deontic classification of clauses

In contrast, REL licenses are only intended to express rights of the licensee subject to certain conditions (what licensee may do, only one of the four categories described before). In fact, REL licenses can only express a subset of the typical rights and conditions given in practical contracts.

2.3. Electronic Contracts Requirements

The first question is to decide whether all the kinds of clauses will be represented or not. The answer comes from the double requirement that we impose to an electronic contract. In one hand, the electronic contract should keep the same information conveyed in the narrative contract, so that in case of a legal dispute, a reference agreed document exists. And in the other hand, some of the information within the contract should be represented in a machine-understandable way. Our analysis leads to the necessary distinction of those clauses whose enforcement and control can be put in hands of a computer ("license clauses"), and those "lawyer clauses" whose interpretation would always be left to the human intellect.

The "license clauses" will be then semantically represented in the electronic contract, while the "lawyer" clauses will be present as a REL r:otherInfo element only for notarial purposes.

In this context where some clauses are to be digitally managed, two different scenarios could be considered. In the first one, a neutral party would keep the electronic contract and would act as a central system supervising the obligations of the two parties (as in [4]), and in the second, one of the parties would host an authorisation system to check that the other party´s restrictions are satisfied. This second case is the most practical, because usually one of the parts is the one that bears most of the obligations and conditions. Therefore in this paper we will only consider this second case, which matches essentially with REL languages scope (expressing rights of the licensee) although REL covers it in an insufficient way. The next section describes breifly the MPEG-21 REL and shows how to make it useful for representing electronic contracts.

3. The MPEG-21 Rights Expression Language

Right Expression Languages (RELs) are languages devised to express conditions of use of digital content. They have been proposed to describe licenses governing digital content. Part 5 of the MPEG-21 standard [1] defines a REL that is based on the eXtensible rights Markup Language (XrML) [11].

MPEG-21 REL defines (in its core, standard and multimedia extension) a set of 18 rights, and 24 conditions. It has, additionally, other extensions whose elements can be taken. The MPEG-21 REL can be extended to support new business models defining extensions. The extensions mechanism that MPEG-21 REL specifies allows the addition of new elements to address the requirements of a new application domain Currently, MPEG-21 REL standard specification has four extensions: multimedia, standard, multimedia extension one and multimedia extension two. The standard extension defines terms to extend the usability of the core schema; essentially it defines conditions that restrict the use of the content, for example in the interval of time, number of times that it can be used, the fees that must be paid, the territory, etc. The multimedia extension expands the core schema by specifying terms that relate to digital works. Specifically describes rights, conditions and metadata for digital works, which includes rights as modify, enlarge, reduce, move, adapt, play, print, execute, etc.

Resources as Digital Item Resources and some resource attribute conditions, Digital Item conditions, security and transactional conditions.

On the other hand, two profiles have been specified and included in this part of the standard as amendments to MPEG-21 REL standard. The first one, so-called Mobile And optical Media (MAM) profile [15] addresses the needs of the mobile and optical media domains. Moreover, it facilitates the interoperability with OMA DRM REL v2 [16]. In order to support the requirements of this profile, the "multimedia extension one" was defined with new rights and conditions for the pre-recorded optical media and mobile domain. Then, this profile consists of a subset of the elements defined in the core and in the multimedia and standard extensions and the rights and conditions defined in the multimedia extension one.

The second one, the Dissemination and Capture (DAC) profile [17] was designed to be able to represent the concept of the OMA DRM v2.0 Extensions for Broadcast Support and to facilitate the interoperability with the TV-Anytime Rights Management and Protection information [18]. This profile consists of a subset of the elements defined in the core and in the multimedia, standard and multimedia one extensions and the rights and conditions defined in the multimedia extension two.

4 Analysis of contracts for its modeling in licenses

4.1 Analysis procedure

Independently from what MPEG-21 REL offers, an analysis of the narrative contracts was done, listing the basic rights and conditions that should be considered.

This modeling task was carried out with the supervision of the interested parties, and a representative sample set of 20 real narrative (paper) contracts was considered. These included contracts from diverse material (audio, video, images) binding different parties (traditional and internet distributors, producers, etc.), even in different languages. Contracts presented varied forms, with different structure and different number and size of clauses.

First of all, a preprocessing was performed to rearrange the long clauses into paragraphs with a single idea by splitting the long ones.

Then, clauses were classified into one of the four possible entries in Table 1, and only clauses for the licensee were considered. Rare clauses that appear seldom were dropped from the analysis, and the rest, were clustered into semantically equivalent terms.

4.2 Results of the analysis

The next list enumerates the most common rights expressed in contracts:

To reproduce: to authorize the act of reproduction of content in any manner or form. Reproduction covers all methods of rendering, for instance drawing, lithography, offset and other printing processes, photocopying, recording etc..

To download: to copy data (usually an entire computer file) from a main source to a peripheral device.

To upload: to transfer data from a peripheral computer or device to a central computer.

To make available: the "posting" or storage of material or information on a computer or server connected to the World-Wide-Web or the connection of a computer containing material or information for access using the Internet or an intranet.

To publicly perform: to present or execute a work in a place open to the public or at a place where a substantial number of persons outside of a normal circle of a family and its social acquaintances are gathered.

To exhibit: to show outwardly.

To transmit: to send data over a communications line

To broadcast: to send out or communicate, especially by radio or television

To copy: to manipulate the licensed content in order to produce a new digital object whose characteristics are the same as the original one and which is autonomous from the latter

To publish: to prepare and issue certain material for public distribution or sale

To print: to produce something in printed form by means of a printing press or other reproduction process.

To record: To register sound or images in permanent form by mechanical, electrical or electronic means for reproduction

To modify: to change in form or character

To translate: to render in another language

To dub: to insert a new soundtrack, often a synchronized translation of the original dialogue, into a film.

To adapt/edit: to make suitable to or fit for a specific use or situation

To convert: to change a content into another format

To transcode: the direct digital-to-digital conversion from one codification to another

To remix: to recombine audio tracks or channels from a recording, producing a new or modified audio recording

To distribute: to supply contents to retailers

To sell: to exchange or deliver for money or its equivalent

To advertise: to use an image or extract of the resource with promotional means.

To lease: to grant the right of possession and use of a content for a specified period in exchange for payments

To synchronize: to cause soundtrack and action to match exactly in a film

To license: to grant a license to or for.

To sub-license: to transfer the right to license

To promote: to attempt to sell or popularize by means of publicity

And the next list includes the most common conditions set upon the execution of the rights. The conditions of the contract are more shaped by the negotiation between the parties and therefore drawing a complete list is difficult .

Term: the period of time during which the rights of a contract will be carried out.

Territory: the area where the rights granted might be performed

Exclusivity: this condition regulates if one party grants another party sole rights

Fee: The clause (or the clauses) which regulate this aspect are aimed at disciplining the remuneration of the licensor. Basically the analysis of the contracts has put in evidence that in most of the cases a payment of the remuneration by means of royalties is agreed. It has been already pointed out that the method of this analysis is examining the contracts and emphasizing their characteristics. Nevertheless it should be noted that in general there are two main methods adopted by the parties in order to set the remuneration of the licensor: the payment of royalties and the payment of the royalties coupled with the payment of a certain sum as an advance. It should be noted as well that the new business models are presenting more and more options for the remuneration of the licensor.

Reporting: this condition is strictly linked to the need of the licensor of monitoring the use of the content by the licensee in order to set the amount of the remuneration.

ValidityPeriod: Condition to stablish dynamically the period of time during which the rights can be executed. As a difference from "term", the period is not fixed in advance but is determined by an external event.

All the previous right and conditions should be enough to represent a contract under the form of an electronic licence.

4.4 Extensions for the MPEG-21 REL

In the analysis performed, most of the common rights and conditions expressed in contracts were identified. Their semantics were then analyzed, and a direct matching of concepts was stablished between the list of rights and conditions in the previous section and the rights and conditions provided by MPEG-21 REL and its extensions. Those with no clear correspondence, were attributed to new elements to be added in an extension.

Table 2 shows this mapping. Note, that the prefixes used for these rights and conditions are the following: "r" for the MPEG-21 REL core, "mx" for the multimedia extension, "sx" for the standard extension, "m1x" for the multimedia extension one, and "m2x" for the multimedia extension two. Nevertheless, some of the rights and conditions commonly used in contracts are not defined in MPEG-21 REL. Then, we propose the definition of the correspondent right or condition. Then, these elements will be part of a new extension for the MPEG-21 REL. The prefix for the elements of the contracts extension is "axm".

CONTRACT	MPEG-21 REL
To reproduce	axm:reproduce
To download	m1x:governedCopy
To upload	m1x:governedMove
To make available	r:issue
To (publicly) perform	axm:perform
To exhibit	axm:perform
To transmit	axm:transmit
To broadcast	axm:broadcast
To copy	m1x:governedCopy
To publish	axm:publish
To print	mx:print
To record	axm:record
To modify	mx:modify
To translate	axm:translate
To dub	axm:dub
To adapt/edit	mx:adapt
To convert	mx:adapt or mx:modify
To transcode	mx:adapt or mx:modify
To remix	axm:remix
To distribute	r: issue
To sell	r:issue (sx:FeeFlat | sx:FeeMetered | sx:FeePerInterval | sx:FeePerUse | sx:FeePerUsePrePay)
To lease	Right + Payment Condition + valityInterval
To advertise	r: obtain axm: publicize

To synchronize	axm:synchronize
To license	r: issue r:delegationControl
To sub-license	r: issue r:delegationControl
To promote	axm:promote

Table 2. Contracts elements and MPEG-21 REL mapping

Once identified the rights and conditions commonly used in contracts that are not defined in the MPEG-21 REL, we propose the definition of an extension for this REL. Figure 1 shows the contracts extension.

Figure 1. Extension for the MPEG-21 REL

Table 2 lists 27 verbs extracted from the contracts. This number is arbitrary, as the number of different verbs actually found in the contracts is much higher. A trade-off in the number of defined terms must be reached, and the satisfaction of the parties with the result will be the score of the translation. Thus, if both parties accepted replacing the text in the narrative contract with the precisely-defined REL term, the system we propose would be acceptable. The participation of the interested parties in the audiovisual parties (with legal background) is essential in this or any other technical project with the same goal, and therefore this has been assumed in the development of this model.

4.5 Examples of translation

To illustrate the previous ideas, we show real clauses extracted from different contracts.

Contract 1: We find in one of the contracts the clause:

"Licensor grants to Licensee the exclusive right, privilege and license, […] throughout the Territory of the People's Republic of China".

This would be encoded as follows:

```
<sx:territory>
    <sx:location>
        <sx:country>iso:CN</sx:country>
    </sx:location>
</sx:territory>
```

Given that the namespaces are stablished as `xmlns:sx="urn:mpeg:mpeg21:2003:01-REL-SX-NS"`and `xmlns:iso="urn:mpeg:mpeg21:2003:01-REL-SX-NS:country"`

Contract 2. In this second contract, we find a clause in italian language:

"Per la durata del presente accordo, la LICENZIANTE conferisce a X che accetta, il diritto esclusivo di […] nei territori di: ITALIA – CITTA´ DEL VATICANO – REPUBBLICA DI SAN MARINO"

What would be codified in the same but using the elements: `iso:SM`, `iso:VA`, and `iso:IT`. We note that regardless the human language and on despite of using different words, the meaning is the same and the codification into the electronic license is valid. Not always the mapping is so clear, and sometimes the territories are given implicitly. For example:

Contract 3.

The territory in which Licensor may exercise each and all of the rights granted herein shall be the territory of North, Central and South America ("Territory")

In this case, there should come an enumeration with all the countries of the american continent. This interpretation of the contract cannot be easily performed by a computer in an automated way, but once translated into the digital license language, computer could easily enforce the fulfilment of the condition.

The three previous contracts referred to three different kinds of audivisual material, the first contract was dealing with ringback tones, the second with television material and the third with motion pictures. Note the abstraction that the REL makes, applying the

same concept to different resources and still keeping the consistence.

And once described the desired digital license we would like to obtain from the narrative contract, the next section proposes a method for implementing this.

5. A Semi Automated System

5.1 Description

In practice, there are many narrative contracts that are still in force or are model for future contracts. They should be mapped into their corresponding electronic version with the tags described in the previous section, what might be a dull task. The kind of persons involved in this translation may come from a legal background, and the task may be highly repetitive, so having a user-friendly computer guided system to process the contracts would be rather helpful.

An ideal automatic translation system would parse a narrative contract and without human supervision would extract the electronic version from it. However, implementing this system is far beyond the state of the art. Instead, what is proposed here is a semi-automated or guided process, where the responsible of the narrative contract can easily extract the electronic version with the support of a computer. So, the user introduces the contract in a text format, follows a computer wizard, and finally obtains a final electronic license.

A first prototype of this system has been implemented under the form of a web application. An stable version is still under development for the FP6 Axmedis Integrated Project [19].

4.2 Implementation

The program operates in two steps:

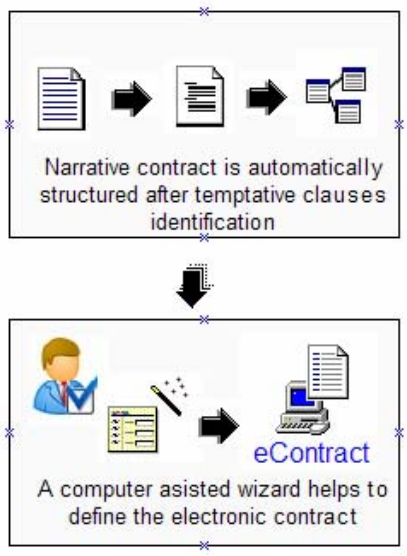

Figure 2. From narrative contracts to electronic licenses

5.2.1. From text contracts to tagged contracts. First, the application converts the text file into an intermediate contract descriptor file. All the sentences in the contracts are statistically analyzed, and those that are likely to belong to one of the given set of clauses, are automatically pre-classified in a new tagged file.

This file is structured as an XML file to be the input of the next stage, and could follow an eContracts Schema [14], containing the following elements:

- Metadata (given as DublinCore elements [20])
- Title (and subtitle)
- Contract-front (date, parties)
- Body (set of clauses, with a temptative classification attending at their kind)
- Back (signatures)
- Attachments (if any)

In order to identify these sections and to give a primitive classification of the clauses, a statistical analysis is done.

This analysis bases its decision in a preloaded database, where each of the considered rights and conditions is associated to a set of typical English keywords, keywords that when analyzing the particular contract will be seeked. For example, the "territory" clauses, usually include terms such as "country", "territory", "region" or "world" etc. Each of these words receives a ponderation, and when analyzing the text contract, an optimal decision will be taken.

5.2.2. From tagged contracts to electronic licenses The mere identification of the parts in the tagged contract, either in eContracts style or any other, is already an important step that would justify by itself

the process of conversion from plain text files to the XML document. It allows a better organized storage of the documents in a contracts database and facilitates their management.

But in order to allow the automatic enforcement of the contracts, a step further must be done and some clause meanings have to be accessible by the computer. Hence a MPEG-21 REL license is generated.

This is done in a guided process, where the user is asked to fill in some forms. The web application will offer sequentially a temptative interpretation of the clauses, that the user will have to confirm or modify the proposed MPEG-21 REL term. Both the rights and the conditions listed in Figure 1 are supported. While this schema works well with some conditions (fee, territory and date), where the vocabulary is rather closed, in other kind of clauses the system may fail to provide a valid suggestion and the user would have to introduce entirely the details.

6. Conclusions and Future work

Although classification of contract clauses and elaborating a good electronic representation as licenses of narrative contracts is by itself an interesting task, electronic contracts and licenses should show their fully potential in the context of practical applications for the audiovisual market. Future work precisely should focus in the development of applications that make use of the electronic contracts.

The process of making the electronic licenses has much to improve too. Natural language techniques could be used for a better automatic classification of the clauses, and for a better extraction of the data. The user then would find the forms already filled with the right information and should require only validating the computer proposal.

This tagged version of the contract can be improved now. Since the new OASIS eContracts standard for structuring electronic contracts has been fixed (it was not present earlier), its adoption for the tagged contract described in section 5.2.1 will foster interoperability with other systems. And given the broad nature of the parties possibly interested in this contract representation format, interoperability is also an implicit goal which parallels with the aim of the Digital Media Project [21], with whom a possible relationship should not be discarded.

The results of the work presented in this paper should be indeed evaluated in the course of the practical applications, in two levels of operation; as structured documents in tagged contracts, and as semantically meaningful documents in the electronic license. Once the experimental stage is overcome the tool should be properly tuned and graphically prepared to cope with general non-experienced users, from which the most valuable feedback has to be extracted. A methodology for carrying out this feedback and evaluation would be desirable and a possible future task.

7. Acknowledgements

This work has been partly supported by the Spanish administration (DRM MM project, TSI2005-05277) and AXMEDIS, a European Integrated Project, funded under the European Commission IST FP6 program. A special mention should be done to the Associazione dei Fonografici Italiani (AFI) [22] for their work in contract analysis used in this paper.

8. References

[1] ISO/IEC. ISO/IEC IS 21000-5 – Information technology – Multimedia Framework (MPEG-21) – Part 5: Rights Expression Language

[2] Ianella, R., Guth, S., Open Digital Rights Language (ODRL), http://odrl.net/2.0/WD-ODRL-Model.html

[3] P.R. Krishna, K. Karlapaplem, A.R. Dani, "From Contracts to E-Contracts: Modeling and Enactment". *Information Technology and Management*, 6, 363-387, 2005. Springer Science, October 2005

[4] M. Morciniec, M. Salle and B. Monahan, "Towards Regulating Electronic Communities with Contracts", *2nd Workshop on Norms and Institutions in Multi-Agent Systems, 5th International Conference on Autonomous Agents.* Montreal, Canada, 2001

[5] S. Guth, G. Neumann, M. Strembeck: "Toward a Conceptual Framework for Digital Contract Composition and Fulfillment", *Proceedings of the International Workshop for Technology, Economy, Social and Legal Aspects of Virtual Goods*, Ilmenau, Germany, May, 2003

[6] Kobryn, C., Atkinson, C., Milosevic, Z., "Electronic Contracting with COSMOS - How to Establish, Negotiate and Execute Electronic Contracts on the Internet". *2nd Int. Enter-prise Distributed Object Computing Workshop* (EDOC '98)

[7] V. Kabilan, P. Johannesson: "Semantic Representation of Contract Knowledge using Multi Tier Ontology". *Semantic Web and Databases* 2003: 395-414

[8] S. Llorente, J. Delgado, E. Rodríguez, R. Barrio, I. Longo, F. Bixio, "Generation of Standardised Rights Expressions from Contracts: An Ontology Approach?", *OTM Workshops 2005*, LNCS 3762

[9] A. Paschke, M. Bichler, J. Dietrich: "ContractLog: An Approach to Rule Based Monitoring and Execution of Service Level Agreements". *RuleML* 2005: 209-217

[10] Content Reference Forum. CEL: Contract Expression Language.
http://www.crforum.org/candidate/CELWG002_celspec.doc

[11] XrML eXtensible rights Markup Language, The Digital Rights Language for Trusted Content and Services, http://www.xrml.org/
[12] Business Collaboration Framework (BCF), the Techniques and Methodologies Group (TMG), United Nations for Trade Facilitation and Electronic Business (UN/CEFACT), http://www.unbcf.org
[13] Organization for the Advancement of Structured Information Standards
http://www.oasis-open.org/home/index.php
[14] OASIS LegalXML eContracts Version 1.0 Committee Specification, http://docs.oasis-open.org /legalxml-econtracts /CS01 /legalxml-econtracts-specification-1.0.pdf
[15] Wang, X., Delgado, J., Barlas C. "ISO/IEC 21000-5/FDAM 1 Rights Expression Language: the MAM profile". ISO/IEC JTC 1/SC 29/WG 11/N8342
Klagenfurt, Austria, July 2006
[16] "Digital Rights Management", Open Mobile AllianceTM , OMA-DRM-DRM-V2, December 2004, http://www.openmobilealliance.org/
[17] Chiariglione, F., Kim, T., Wang, X. "ISO/IEC 21000-5/FDAM 2 Rights Expression Language: the DAC profile". ISO/IEC JTC 1/SC 29/WG 11/N8812
Marrakech, MA, February 2007
[18] TV-Anytime Forum, http://www.tv-anytime.org/
[19] FP6 Axmedis Project. http://www.axmedis.org
[20] DublinCore Metadata Initiative, http://dublincore.org/
[21] The Digital Media Project, http://www.dmpf.org
[22] Associazione dei Fonografici Italiani,
http://www.afi.mi.it/