

<in*d*ecs>rdd White Paper

A standard Rights Data Dictionary

May 2002

Introduction

<indecs>rdd is a consortium based initiative to develop a rights data dictionary. Its purpose is to support the implementation of a rights language for secure exchange of intellectual property on networks. The initiative, based on the original <indecs> analysis (1998-2000: www.indecs.org), has been underway since mid-2001. At that time a consortium of rights holder representatives and providers of services funded the work and engaged the services of Rightscom to lead the technical work of the project.

The <indecs> analysis was a reference model. In order to implement the reference model, a practical reference implementation was required. In April 2001 the International DOI Foundation (IDF) funded a feasibility study for the development of a Rights Data Dictionary (RDD), a common dictionary or vocabulary for intellectual property rights to be based on the <indecs> framework. The aim was to propose a consortium which would fund the development of standard rights terms to enable the exchange of key information between content industries for eCommerce trading of intellectual property rights. <indecs>2rdd was the next phase in <indecs> development. Because rights metadata is inseparable from other metadata, and because the <indecs> framework specifies a general metadata framework, the work done in developing <indecs>2rdd also deepened and expanded the original <indecs> framework, building on it.

The first stage of the initiative ended with the submission of the <indecs>rdd fundamental design to MPEG¹ for the purposes of standardisation: in December 2001 the initial <indecs>rdd design specification was selected as the baseline for the MPEG-21Part 6 Standard for a Rights Data Dictionary.

The second stage of the initiative followed this acceptance, and the <indecs>rdd consortium continued, with a revised set of partners, with two objectives:

The primary objective is to complete the MPEG-21000 Part 6 standard for a rights data dictionary. To achieve this, the consortium has retained Rightscom to work in collaboration with MPEG experts to create and edit the standard. It is expected that a Committee Draft will be produced in October 2002, at the 62nd meeting to be held in Shanghai, China. The Committee Draft will be balloted by MPEG national bodies and comments returned in time for discussion at the 63rd meeting, in March 2003. The International Standard could then be declared in the autumn of 2003.

The second objective, which both supports and is dependent on the first objective, is the development, on behalf of the consortium, of a version of the rights data dictionary, complete with an extensible database design and

¹ (ISO/IEC JTC1/SC29/ WG11)

interface, which could become the basis of an operational system. This objective is dependent upon the MPEG process, in that the database and dictionary must be conformant with the MPEG standard. Furthermore it also supports the first objective by providing proof that the standard can be implemented fully and that there is nothing in the standard that cannot be completely realised in operational terms. Some consortium members will use the developing dictionary in early implementations as a practical demonstration of their confidence in the work.

The need for a standard rights data dictionary

“Rights management” has to work in a computer environment. The huge amount of digital content now being traded – legally and illegally – requires an infrastructure for rights. What does “rights management” mean in this context?

The *terms* used in various “rights expressions” which mediate the use of digital items – ownership statements, licenses, permissions, offers, requests and agreements – need to be *unambiguously* understood by computers. Together these terms are often called *rights metadata*.

For instance, if a license agreement states that a commercial *consumer* must *pay* a particular fee to *copy, play* and *keep* a particular *format* of a *digital file* in a particular time and place, and that a *student* may do the same for a *reduced price*, all those terms must be interpreted by a computer or user to mean what is intended by the licensor.

To achieve such a level of unambiguous interpretation, there must be a common *data dictionary* of terms involved in rights. This is a common requirement in computing, but in the area of rights management there are three problems which make it especially challenging.

Three Problems

First, rights are complex. Rights metadata can quickly become much more complicated than the simple license example given above: all kinds of media, content and usage might be involved, including rights in underlying abstract works; ownership of rights often changes over time. A rights data dictionary must be capable of supporting the simplest through to most complex of rights expressions.

Second, rights expressions will be mixed with other types of information. Agreements, offers and licenses may include any terminology taken from descriptive, legal or financial systems. A rights data dictionary must be broad enough to embrace terms from any other kind of metadata which might occur in a rights expression.

Third, many dictionaries are already in use. Different market sectors, individual companies and organizations may have their own working dictionaries and schemes (often called different namespaces). Some deal with rights, some don't. Many groups will not want, or be able, to change to a new dictionary, or use a new one alongside the terms from their own namespace. Yet because these groups are now all co-operating in common multimedia areas, some way of connecting them is essential. A rights data dictionary must be allow the use of terms from existing and future namespaces.

The solution: <indec> rdd

The architecture for <indec> rdd has been developed over a number of years to cope with just these problems of complexity and interoperability. A call by MPEG early in 2001 for proposals for an RDD standard provided the impetus for turning this architecture into a practical tool in the form of a distributable data dictionary with a range of powerful features. It combines the main elements of a data dictionary, a multi-lingual dictionary, an ontology and a thesaurus.

<indec> rdd is well suited to this task because:

It has a powerful conceptual base. <indec> rdd is based on a strong and mature underlying data model (the "Context Model") in which *verbs* are the starting point for all the most important definitions. This model provides a core of several hundred terms to which any number of others may be added in a systematic way.

It is highly structured. Every <indec> rdd term has a unique identifier and a "Genealogy" which defines precisely and logically how it relates to others. Because the underlying model is very rich, it can accurately describe very complex relationships between terms.

It is inclusive. Any terms from other dictionaries can be added to <indec> rdd (by assigning a unique identifier and a genealogy). Other terms are not just "extensions" or "mapped" words, they become an integral part of <indec> rdd itself.

It is highly granular. <indec> rdd can support terms at any level of detail, fragmentation or versioning required by users.

Users can "mix and match" terms. Because any "mapped" scheme is part of <indec> rdd, terms from different namespaces can be combined to form rights expressions without loss of meaning. Each namespace may have and use its own set of <indec> rdd terms to create its own rights expressions.

It supports "transformations". <indec> rdd provides the *semantic* tools needed to translate terms from one scheme to another in a highly automated way. This is critical to allow different metadata schemes to co-exist in the multimedia environment. Software applications are still required as well, but <indec> rdd provides all the underlying semantic relationships.

It is legally neutral. <indec> rdd does **not** define legal terms. It can be used to make rights expressions which draw on any existing legal definition, or none.

It is business-model neutral. <indec> rdd terms can be used to describe any situation in which any kind of rights are owned, managed, protected or

used, at any point in the life of content from origination to "end use" or archiving.

It is not a Rights Expression Language (REL). A data dictionary is not an expression language (such as XrML, now adopted as baseline technology for the MPEG-21 REL standard). An REL deals with the way in which terms are *expressed* in computer language; the dictionary *defines* the terms. An REL will use terms defined in an RDD.

It has uses beyond rights. Because of its generalized model, <indecs>rdd can provide a comprehensive basis for metadata expressions and schemes for purposes other than rights – such as resource description, workflow management and event reporting. It could be used as a tool for the deployment of semantic based web services.

How <indecs>rdd will be used

<indecs>rdd is a tool which will be used in an automated way (often invisibly) to help to create, transform and interpret rights expressions.

It will:

Provide a ready-made standard terminology for rights management.

Organizations needing to create rights expressions, or to enhance their existing metadata schemes, will be able to use <indecs>rdd as a source for terminology. Apart from providing a structured basis for metadata selection, it ensures interoperability with other compliant schemes.

Be available in a variety of forms. The dictionary will grow constantly as other schemes are mapped, and so (as with "virus checking" software) regular updates will be an essential component.

Support application software at all points in the "content chain".

<indecs>rdd will be available to support the making, transforming and interpreting of rights expressions from origination to end use.

Who supports <indecs>rdd?

The <indecs> consortium represents major groups of rights owners and ancillary service providers. Currently its members are:

Enpia Systems Ltd.

International DOI Foundation

International Federation of the Phonographic Industry (IFPI)

MMG Ltd.

<indecs>rdd White Paper May 2002

Motion Picture Association (MPA)

Recording Industry Association of America (RIAA)

The <indecs>rdd project is managed by Rightscom (www.rightscom.com)

For further information please contact:

Chris Barlas
Rightscom
10 Leake Street
London SE1 7NN, UK

+44 (0)20 7620 4433

chris.barlas@rightscom.com