Open Digital Rights Language (ODRL)

DR RENATO IANELLA

Note: this paper has been updated for December 2007

Let me start with a few words about Digital Rights Management (DRM). As usual, it was mentioned in other talks in the negative, which is fair because DRM does have some negative aspects about it. But I want to give you a different view from the DRM world. Then I will look at the Open Digital Rights Language (ODRL) Initiative. I will then look at the Creative Commons' semantics in more detail, how we mapped them to ODRL and some of the issues that we found when we were doing the mapping exercises that we feel are quite significant and should be raised. Finally, briefly, I will look at the potential to do a similar thing for the AEShareNet licences.

Creative Commons licences are represented in three ways: there is the legal code, the human code and then there is the machine code. All three are very critical to the overall licences, but what I will be presenting here is more aimed at the machine code. I will ask: how have the licences been represented at the computer level and how can they best be represented?

DRM covers two main areas. There is the information about the rights - the rights information management - and that is about who the rights holders are, what the licences are, what the royalty payments are, etc. Then there are is the enforcement/security side, or the technical protection measures, including the trusted environments. This is usually that area that gets DRM bad press as it is squarely at the consumer end. The consumer sees the way the content is encrypted and the way that limits the end user experience or changes the way the end user has to interact with that content. A lot of current DRM systems really do just focus on the security side and do not care about rights information management.

There are positive examples of DRM working, such as Apple's iTunes/iPod service. Most of the consumers who buy and download songs to their iPods have no idea that DRM is in there because it is well hidden, which it should be. And it still allows the consumer to do what they normally expect to do with their music, which is just play it an unlimited amount of times and also, in some cases, to make copies for a fixed number of times to different devices.

127

The technical view of DRM also needs to be balanced with the social, legal and business sides. The DRM value chain needs to support both of the two DRM areas as the rights management information normally has to travel from the beginning to the end of the value chain (ie from when content is created to when it is being used) and at the end, we need to have the rights information there. The enforcement is usually at the consumer end, so it is downstream, the last thing that happens. Usually, the content is encrypted or somehow encoded so that only a particular consumer or device can consume it.

The key here is in the rights management information metadata that is being captured in what is now called 'Rights Expression Languages' (REL), a new sub-discipline, if you like, of DRM. In terms of standards, there are basically two standard bodies that deal with DRM at the international level. There are others, but the main ones are the Open Mobile Alliance, which is the mobile sector, and then there is the MPEG-21 standard, which is the audio/video sector. In MPEG-21, Parts 4, 5 and 6 deal with DRM. There is at the moment up to fourteen parts, but those three deal with DRM. Since 2000, we have seen a bit of a standards "war". There was a battle between two rights languages, XrML and ODRL, and the two different standards bodies chose two different languages. There is a lot of politics behind that, which makes life interesting, but it basically came down to the typical "Microsoft versus The Rest of the World" battle; Microsoft owning XrML and "The Rest of the World" not wanting that technology in their standards. To this date, the war is not over. In fact it is probably hotting up at the moment because there are also a lot of DRM patent claims being fought over in this area. This is going to make life very difficult for implementers of DRM systems and devices because it will make it uncertain as to what your liabilities are.

In early January 2005 the MPEG Licensing Authority issued a press release stating the terms and conditions for licensing the Open Mobile Alliance (OMA) DRM specifications. Interestingly, we have one standards group telling the other standards groups how much they are going to have to pay to implement their own DRM standard. It is fun and games in that area. ODRL and XrML are two rights expression languages, which are extensible expression languages. You can express anything you like in them, but they do come with their own dictionaries of common terms.

The scope of RELs is explored in a report which came out of the UK in

2004[42]. It looked at where rights expressions are captured in the entire value chain and it went through these processes:

- Recognition of rights,

- Assertion of rights,

- Expression of rights,

- Dissemination of rights,

- Exposure of rights, and

- Enforcement of rights.

The ODRL Initiative is an initiative that has been running since 2000, originally developed by IPR Systems. They obtained additional partners, like Nokia and RealNetworks, and incorporated their specific rights expression languages into the ODRL language. The ODRL Initiative has an independent governance board that looks over the governance issues and promotes ODRL to larger standards groups. They have had success in OMA, and have also published a World Wide Web Consortium (W3C) Note. They have also submitted ODRL Version 1.1 to National Information Standards Organization (NISO), which is the US standards body.

A number of ODRL working groups are now looking at how to develop the language further. We have one looking at Version 2 and, of course, the Creative Commons Profile working group. We are also looking at GeoSpatial data and, in early 2005, the Dublin Core Joint Working Group was announced to look at how to use Dublin Core and ODRL together. We are also planning a NISO/Library Joint Working Group that will look at joining the needs of the library community with ODRL.

The core model of ODRL is shown in Figure 1. There are three main aspects we look at: 'rights', 'parties' and 'content'. Parties and content can be further exploded into different aspects, as well as the core rights in terms of the permissions, constraints, the conditions and requirements. These are the key aspects to any rights expression language, not just ODRL.

---

[42] Joint Information Systems Committee (JISC), *JISC Digital Rights Management Study* (2004) <http://www.intrallect.com/drm-study/>
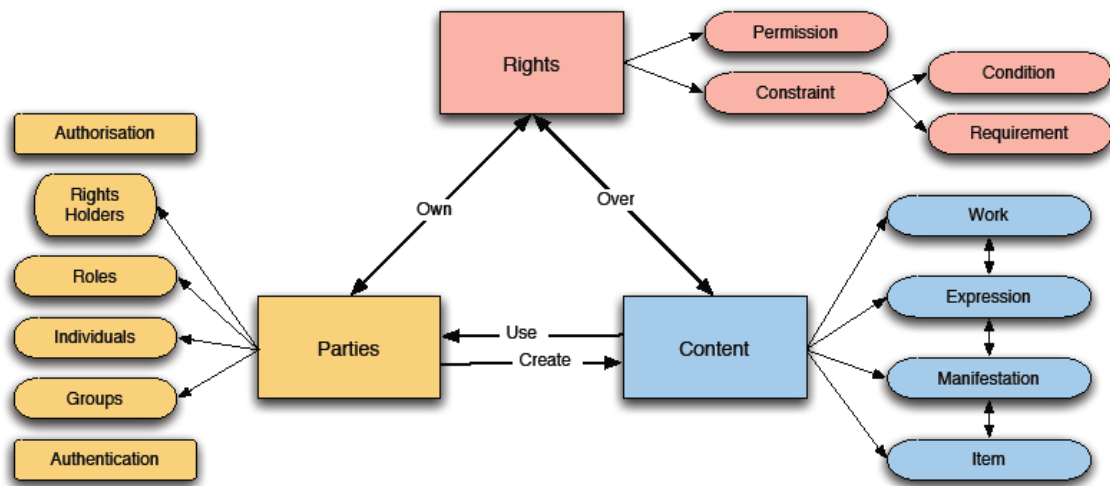
Figure 1

In a commercial example (see Figure 2), we have a famous author or retail store owning some content, having some rights holders, and able to embark upon an agreement to sell the content with a particular constraint (eg a specific country limitation). Each requirement and constraint is optional so you do not have to always have a payment attached, you can simply express the rights in that value chain, for example in the education sector. Then that person or retail store can then make subsequent agreements with other people (eg Joe Consumer) to acquire the content under different types of conditions and constraints (eg print only once).
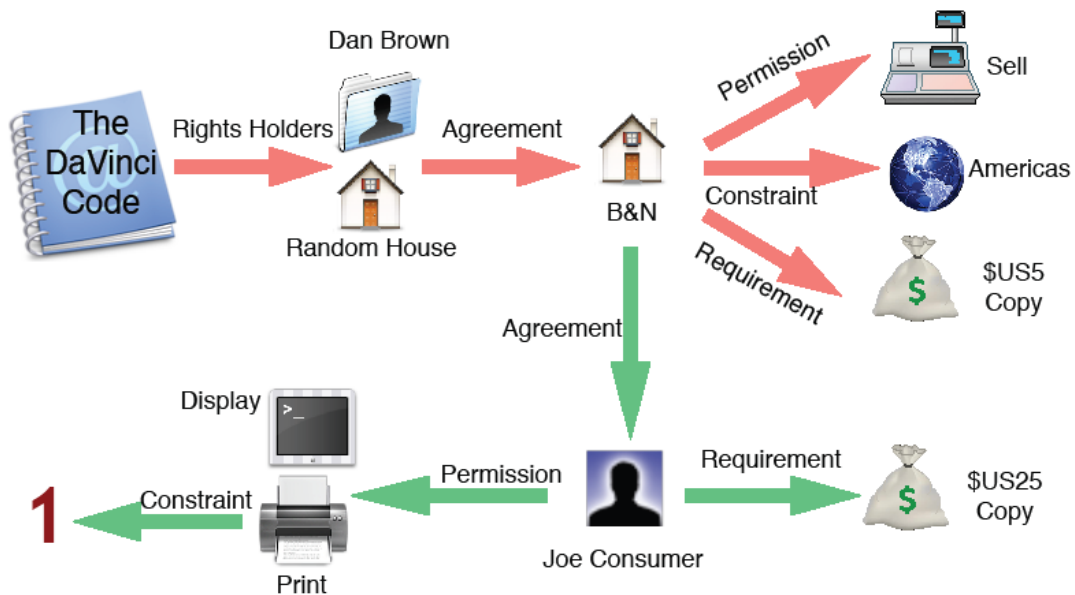
Figure 2

Creative Commons has three aspects to its licences. There are 'permissions', 'prohibitions' and there are 'requirements' – similar to the

ODRL model but not exactly the same. Each of those aspects have a number of fixed values under them and a collection of those makes up a particular Creative Commons Licence. They are technically expressed in RDF/XML, which is another issue for RELs, because the major RELs are expressed in XML Schema and not RDF/XML. That becomes a bit of an interoperability issue, but something that can be overcome.

The ODRL/Creative Commons Joint Working Group started at the end of 2004 and released the final specification in July 2005[43]. The motivation was to see how to express CC semantics in the ODRL language and the benefit of that is that it allows users to use a more expressive language – so that they can then add additional and key information to those licence terms. The actual core semantics of a CC licence includes:

- three permissions – reproduction, distribution, derivative works

- one prohibition – commercial use, and

- four requirements – notice, attribution, share alike and source code.

A combination of these makes up various CC licences.

To give you an example of some of the issues we found with the CC licences, if you look at 'attribution' for example, 'attribution' says that "credit must be given to the copyright holder and/or author". But in many cases, if you acquire content under this attribution licence, you may not know who the copyright holder is or the author. If you have a music file, or an audio file, it does not tell you who the author and the rights holders are, and how to attribute them whenever you reproduce this content for your own uses. We are still lacking some key information there. We need to be able to specify who are the authors of the content, who are the rights holders and how should you attribute them. Do you pop up a window, or do you write a bit of text on the screen? How do you do that?

Each CC licence has its own unique identity, via a URL, and is made up of a set of permissions, constraints and requirements. When we map these to ODRL, the permissions were the same as ODRL, requirements are the same, but prohibitions were not in our model because in ODRL, and other rights expression languages, we have the concept that whatever is explicit in the licence is what is only allowed. If you do not allow something, then you do not put it in licence. If you do not allow commercial use, then you

---

[43] 'Creative Commons Profile' *Open Digital Rights Language (ODRL) Initiative* (July 2005) <http://odrl.net/Profiles/CC/SPEC.html>

do not put it in the licence. It is very simple. Whatever is in the licence is what you are allowed to do.

What we had to do is map the CC prohibitions to ODRL constraints and then change them from the negative to the positive. We can have a constraint of commercial use, which is the same as a prohibition of non-commercial use, and that is what we had to do in the ODR/CC profile. The CC semantics have much broader concepts than some ODRL terms. They have terms like 'reproduction' whereas ODRL has terms like 'print', 'display', 'play', 'execute' – very specific terms that are obviously meant for a machine to interpret and manage. We decided that it was probably not a good idea to map to those four low-level terms because reproduction could include more than those four. We created new semantics for 'reproduction' and the other broader CC concepts. The same for the requirements, but we did have attribution as part of the ODRL data dictionary so we used that directly. See Figure 3 for the final mapping.

CC Licence	ODRL Permission	ODRL Constraint	ODRL Requirement
Attribution	Reproduction Distribution Derivative Works		Notice Attribution
Attri-NoDerivs	Reproduction Distribution		Notice Attribution
Attr-NonComm-NoDerivs	Reproduction Distribution	Non Commercial Use	Notice Attribution
Attr-NonComm	Reproduction Distribution Derivative Works	Non Commercial Use	Notice Attribution
Attr-NonComm-Share	Reproduction Distribution Derivative Works	Non Commercial Use	Notice Attribution Share-A-Like
Attr-Share	Reproduction Distribution		Notice Attribution Share-A-Like

Figure 3

Some of the additional features that you can use after we have created ODRL licences are that you can specify who the rights holders are, specify details of attributions, have greater fine-grained control over constraints, such as country or regions. If you want to allow distribution of your content but only within geographical bounds we can specify that.

We can also identify the asset directly as well. Creative Commons' licences do not directly identify the asset; they just assume it is been linked to from

somewhere else. And, of course, we have a much richer set of permissions, constraints and requirements, etc. The other additional benefit is that ODRL has the identity of the person accepting the agreement. It does not have to, but it can allow you to be very specific about who is accepting this agreement. With CC licences, most of those licences are implicit – you just implicitly accept them – versus in ODRL we can make it very explicit. The benefit for that is, as the end consumer, I can then have a transaction that says, 'yes, I have got your content and I have got them under these conditions', so that I can use them for the conditions specified and there is no way that you can then say, later on, "well, I did not, I have changed my mind and I am withdrawing that". It gives the end consumer a bit more confidence that they can use the licences.

We found a few more examples of some of the mapping issues in some of the other licences of Creative Commons. For example, sampling licences allow people to take your work and transform it, for any purpose other than advertising. The problem we found was that there were no new semantics defined in the CC machine code for these licences. There are other prohibitions, like commercial use, but they did not define the semantics for prohibiting advertising, which is clearly part of a licence, but does not appear in the machine semantics. There are a few issues like that we are feeding back to the Creative Commons team to see whether they can update their machine semantics to make it clearer what you can and cannot do in the licences. The same problem exists in the music sharing licence. This licence says it has the same semantics as the "attribution, non-commercial, no derivatives" licence, which says you are free to copy, distribute, display or perform the work. But the music sharing licence says legally you can "download, copy, file, share, trade, distribute and publicly perform it". Trade is part of the licence description and could be misinterpreted in many ways. It could mean selling it – which is against the non-commercial term - so there are some semantics that need to be seriously tightened up.

Another final example is in the CC developing nations licence. They have created new licence semantics because the developing nation licence allows your work to be used royalty free in any nation that is not classified as a high income economy by the World Bank. The semantics include the standard permissions and requirements, and they have added a new semantic called 'high income nation use', which is a prohibition. They have added this extra prohibition in the licence description, but it does not appear in the machine code.

The ODRL/Creative Commons profile will show how you express CC licenses in the ODRL language. As another example, we are looking at potentially doing the same process for the AEShareNet licences - to be able to take their semantics and represent them in machine code. It will be very interesting because we will then be able to mix some of the AEShareNet concepts of vetting and consolidation with some of the Creative Commons concepts of notice and attribution, with some of the additional ODRL semantics. It is a good example of mixing different semantics together for different licence profiles.

There are other aspects of the licences that we need to develop further and ODRL Version 2 is still evolving. But one of the things we always had feedback on was how you support copyright exceptions in these licenses, because most other agreements will exclude the copyright exceptions. In ODRL Version 2 - that we are currently working on at the moment - we will allow people to put in an explicit part of the licence that says that the copyright exceptions from a particular jurisdiction have to be honoured as part of the licence terms. We cannot get into the specific details of what those exceptions are because it depends on the content and the jurisdictional laws, but we will be able to put that in the licence - so that you can make available to people agreements that allow them to preserve the copyright exceptions for use in the traditional areas.

To sum up, we notice that rights expressing languages are, in many cases, too expressive. That is why a number of community profiles are now being developed. Creative Commons is a very good example of that because we need to focus on what is needed by the end consumers and by the content providers and express those licences only. One of the interesting ideas is whether there is a potential to consolidate the ODRL language with the Creative Commons machine language. The more times you create additional machine languages, the more programming is required and software will not interoperate, etc. so this could be a potential to try and consolidate the two languages together.