Applying Digital Rights Management Systems to Privacy Rights Management

Korba, Larry; Kenny, S.

Publisher’s version / Version de l'éditeur:

Journal of Computers and Security, 2002

National Research Council Canada
Institute for Information Technology

Conseil national de recherches Canada
Institut de technologie de l'information

NRC-CNRC

*Applying Digital Rights Management Systems to Privacy Rights Management **

Korba, L., and Kenny, S.
November 2002

* published in Journal of Computers and Security, November 2002. NRC 44955.

Canada

Applying Digital Rights Management Systems to Privacy Rights Management[1]

Steve Kenny
Dutch Data Protection Authority
Den Haag, The Netherlands
Stephen_MH_Kenny@yahoo.com

Larry Korba
National Research Council of Canada
Ottawa, Canada
Larry.Korba@nrc.ca

Disclaimer

The views expressed by the authors of this article are their own and may not necessarily be taken to be those of either the Dutch Data Protection Authority or the National Research Council of Canada.

Abstract

While there are growing concerns about how to manage citizen privacy, currently there are no established technology solutions that meet the privacy needs required in some cases by legislation. In this paper we examine the prospect of adapting systems developed for Digital Rights Management to meet the challenges of Privacy Rights Management. In particular, the goal of this work is the adaptation of DRM technology to produce a privacy management architecture that reflects the requirements of Directive 95/46/ECfor the protection of personal data. This paper first outlines the requirements for management of the personal data of citizens of the European Community. It then describes the changes that would be required to transform a digital rights management system into a system to manage the handling of personal data. The paper concludes with a thorough discussion of the issues and potential of this approach.

1. Introduction

Privacy issues facing developed societies today are made complex because of differing ideologies and policy amongst countries (e.g. the United States and the European Union), the rapid deployment of networked commerce and technology evolution in general. In this context, privacy issues from a technological perspective are complex due to Directive 95/46/EC of the European Parliament and the Council of 24 October 1995. This legislation, hereinafter referred to as the Directive [1], describes the protection of individuals regarding the processing and free movement of their personal data. Many of the provisions of the Directive have the potential to become global, de facto standards for e-business.
In this paper we investigate the potential for applying Digital Rights Management (DRM) technology for Privacy Rights Management (PRM). We define PRM as the management of personal information according to the requirements of the Directive. Regarding drivers for PRM, we do not consider present DRM risks to privacy in an explicit sense, but rather the potential of adapting DRM architectures to manage aspects of the Directive's requirements for handling personal data during the course of network-based activities. We develop this approach as a framework for a broader integration of privacy services, and to uncover pertinent research issues towards the development of new, privacy management architectures.

The rest of this paper is organized as follows. Section 2 offers a brief overview of DRM systems. In Section 3 we sketch an overview of EU data protection infrastructure. Section 4 describes the components of a PRM system as it reconfigures and extends the DRM model. Within a PRM system some mechanism is required to express privacy. In section 5 we provide examples of how this may be done in the context of a DRM system. Section 6 provides a detailed discussion on this analysis.

[1] NRC Paper Number: NRC 44955

2. DRM overview

DRM is a technology originally conceived to facilitate controlled distribution of digital information in an attempt to combat breaches of copyright law [2]. An advantage of DRM technology is that it offers a technique to control and bill for digital content usage, through persistent information protection (PIP). DRM systems essentially provide licensing, plus the ability to change content access, on the fly. This occurs through databases that structure transactions and relationships among rights, works, and parties by way of a number of technologies.

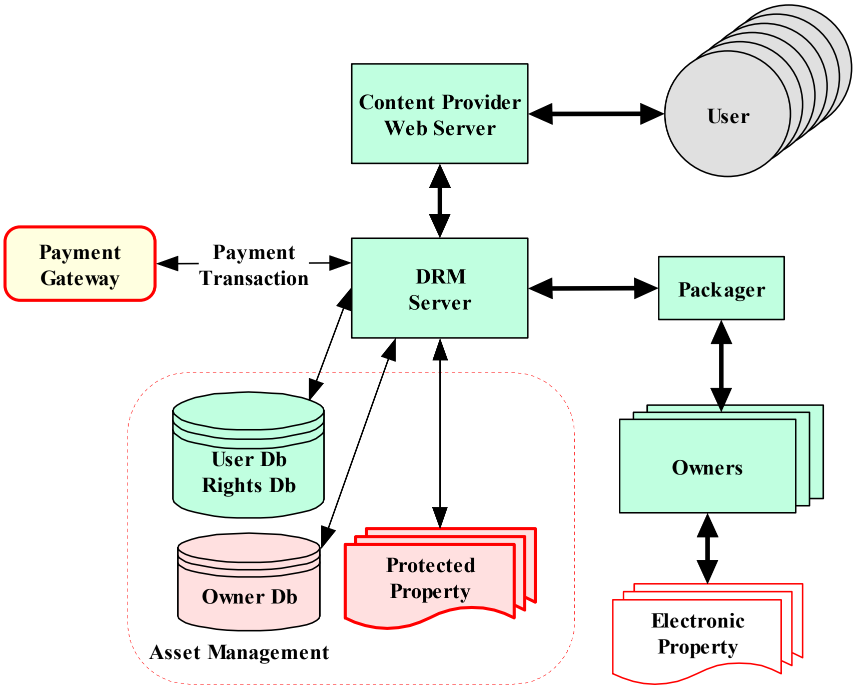

Figure 1. A simplified DRM System

In order to present the concept of privacy rights management, we adopt the Client-server rather than the peer-to-peer architecture for reasons of simplicity. As illustrated in Figure 1, a DRM system operates in the following fashion. An owner or distributor submits its electronic property to a packager that encodes the property into an appropriate format for eventual end use. The packager encrypts the content to guard against unauthorized use, and adds metadata. The metadata not only specifies the content, but also may hold information regarding under what circumstances a user may gain access to the contents. The DRM Server, sometime known as the Rights Fulfilment Server, manages assets stored within various databases. An important concept that forms a foundation for DRM is the separation between the content and the rights for access to the content. Rights describe precisely what a user is allowed to do with the content. Typically some sort of language is used to express those rights (for example: XrML [15], and ODRL [20]). The Rights Management Language implements the business model for the commercial distribution of the content providing details concerning different types of purchase models, use models, etc.

In order to view or play DRM managed content, the user must deploy client software on his computer. This client software handles user authentication and provides secured access to the content. The intention here is to ensure that only those entitled to a file will be able to access it. A challenging element of data protection while in the hands of the data processor is ensuring that the content may not be saved when it is available in the clear for processing.

DRM accommodates several content types and many layers content distribution, ranging from content locking mechanisms (through wrappers incorporating encryption algorithms and access control) to content metering, payment processes and record keeping. DRM architectures support description, trading, protection, monitoring, tracking, transfer and use of digital assets. For instance, DRM technologies can control file access (screen capture, number of views, length of views) and file use (altering, sharing, copying, printing, and saving). Technologies supporting DRM may be contained within the operating system, program software (for example, specialised viewers), or embedded within the actual hardware of some devices [21].

DRM architectures can propagate a variety of business models from controlled internal and B2B content sharing as offered by vendors such as eMeta [3], DigitalOwl [4], dotEncrypt [5], Alchemedia Technologies [6] and Authentica [7] through to entertainment facilitation from Microsoft [8] and hybrid offerings from IBM [9], InterTrust [10], Perimele [11] and Viaquo [12].

3. EU background on data protection

In Europe, the right to privacy is defined as a human right under Article 8 of the 1950 Convention of European Human Rights. That right when it is extended to the protection of European Union Citizens' personal data is expressed in a single legal instrument – the Directive.

The Directive defines a set of rights concerning personal data accruing to individuals and a set of rules for lawful processing on the part of data processors applicable irrespective of sector of application. The implicit principles and constructs of the Directive define the enforcement and representation of privacy. Application of the Directive is termed "data protection". For the purposes of this study, we shall use the term "data protection" interchangeably with "privacy protection", although we are aware that in other contexts the two terms are not necessarily commensurate with each other. In this paper, we define privacy as an individual's ability to express control over how his or her personal information is used.

The Directive applies to almost all sectors of public life. It specifies the data protection rights afforded to citizens, plus the requirements and responsibilities of individuals and organizations defined as "data controllers" and "data processors". This triad structure of entities balances data subject fundamental rights against the legitimate interests of data processors. The national legislature of EU member states must implement the Directive to substantially similar degrees. Implementation encompasses sanctioning national enforcement bodies such as the Dutch Data Protection Authority and the Information Commissioner in the UK, with powers of prosecution. the Directive describes the rights and responsibilities of the data controller, data processor and data subject.

A data subject is a natural person who can be identified by reference to one or more pieces of data related to his physical, physiological, mental, economic, cultural or social identities. Article 26 and Article 2(a) of the Directive may be so interpreted that data associated with an individual in ambiguous ways may be deemed reasonably identifiable, a fact that should not be lost data aggregation strategies applied by both business and governments. Following Article 1 of the Council of Europe Convention 108/81 [13] the fundamental right to data protection falls not to the nationality of the data subject, but as an obligation to a relying party of the data subject. The relying parties are the data controller and the data processor.

The data controller is a legal entity that determines the purpose and means of processing personal data, and is defined as the holder of ultimate accountability as it relates to the correct processing and handling of the data subject. An example of a data controller is the legal entity operating an information system that gathers personal data.

The data processor is the entity that processes personal data on behalf of the data controller, as would be the case for use of outsourced data warehousing for instance.

The Directive's content and enforcement apparatus are to varying degrees mirrored in countries such as Canada, New Zealand, Australia, Norway, Switzerland and Hong Kong. Article 26 of the Directive restricts data transfer outside the EU, with the effect of forcing the Directive to be the *de facto* data privacy standard globally for EU data controllers. US organizations are usually untouched by the Directive as long as they have no data controller legal entities based in the EU. The US does not regulate information privacy, but rather encourages individual self-determination over whether an individual chooses to use data relating to them as a commodity. In some cases, this approach has led to the possibility of widespread aggregation of personal data without anonymity of citizens [23].

As the Directive concerns itself with data processing, one generally must deal with its compliance in terms of the application of information technology as well governance initiatives. One of the difficulties involved in achieving compliance with the Directive is the fact that many of the technologies that may be used as components for an application were developed under privacy-violated assumptions – for instance the absence of minimised data collection in accordance with Recital 28. Often new applications are developed without a careful and skilled analysis or audit of the application to determine whether or not it will meet the requirements of the Directive.

Privacy principles abstracted from the complexities of legislation have been developed to simplify compliance with privacy regulations. Such principles offer a fruitful means for a detailed analysis of how a technology may protect data privacy. Table 1 offers a description of the facilitation principals.

Table 1. This table outlines the privacy facilitation principles and the responsibilities of each entity in a PRM system.

Principle	Description
Reporting the processing	All non-exempt processing must be reported in advance to the National Data Protection Authority.
Transparent processing	The data subject must be able to see who is processing his personal data and for what purpose. The datacontroller must keep track of all processing performed by itself and the data processors and make it available to the user.
Finality & Purpose Limitation	Personal data may only be collected for specific, explicit, legitimate purposes and not further processed in a way that is incompatible with those purposes.
Lawful basis for data processing	Personal data processing must be based on what is legally specified for the type of data involved, which varies depending on the type of personal data.
Data quality	Personal data must be as correct and as accurate as possible. The data controller must allow the citizen to examine and modify all data attributable to that person.
Rights	The data subject has the right to improve his or her personal data as well as the right to raise certain objections regarding the controller's execution of these principles.
Security	Measures are taken to assure secure processing and storage of personal data.
Data processor processing	If data processing is outsourced to data processor, controllability must be arranged
Data traffic outside EU	Exchange of personal data to a country outside the EU is permitted only if that country offers adequate protection, assuring appropriate measures are take place in that locality if possible

We use these principles to consider the appropriateness of adapting technologies currently used for DRM to develop a system for PRM.

4. Privacy Rights Management

As is clear from the description of the facilitation principles in Table 1, there are many demanding requirements placed upon the data controller. In order to examine the possibility of meeting these requirements, we propose the system shown in Figure 2.

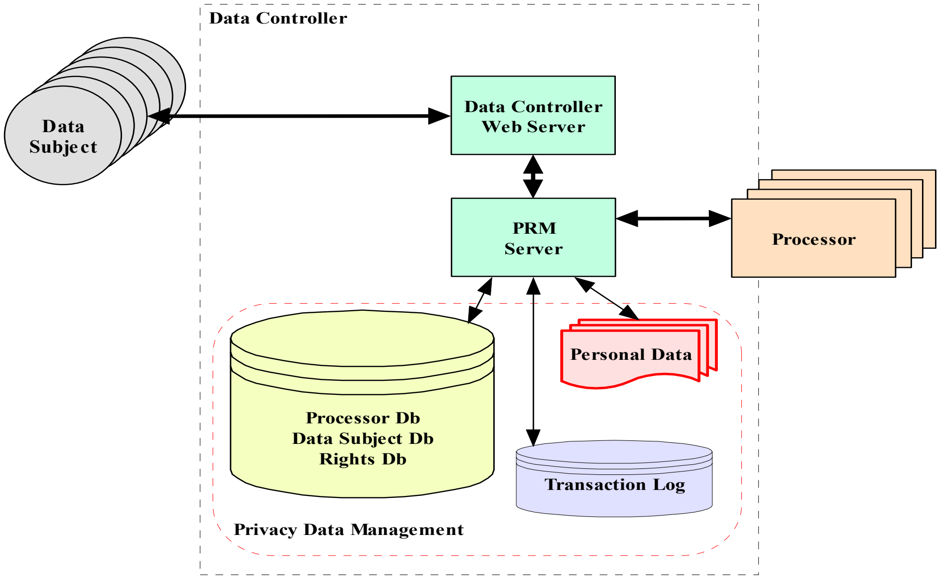

Figure 2. A simplified PRM system.

The participants in the system are the data subject, the data controller and the data processors. The data subject is the originator and *owner* of the personal data managed by PRM. Effectively this arrangement has its boundaries defined by the Directive.

The data controller manages the gathering, storage and processing of data subject data. Data controllers shoulder the ultimate responsibility for proper processing, enforceable through both national data protection authorities and the preservation of data controller reputation. There may be many data processors associated with a PRM system. A data processor may be an element operating under direct control of the data controller, or it may operate as a separate entity, having a contractual arrangement with the data controller. Therefore in PRM there may be many data processors dealing with data from many different data controllers.

From Figure 2, the key components comprising the PRM system managed by data controller are:

- The data controller web server, which is a web server providing the user interface for all of the participants. In a DRM system, after authentication is established, a globally unique identifier (GUID) is applied to user content, enabling ‘profiling’ of the individual. On this step, a PRM system could seek recourse here to security approaches, extant for some time, for providing anonymity. Solutions range from Trusted Third Parties (TTP) to Private Credentials [14]. The latter provides strong anonymity while preventing fraud.

- The PRM server block contains the elements providing base PRM services. Personal data in a PRM system plays a similar role to that of Protected Property in a DRM system (see Table 2). The Data Subject owns the data, and entrusts it to the PRM server wherein it is protected and otherwise managed by the data controller. This relates to the European notion of informational privacy wherein there is, by default, a degree of data subject control over the use of that data. As such, data subject profiles are treated as a data subject system asset. In order to perform its functions, the server block must maintain and use different sets of data. As well, it will manage data exchanges with data processors to meet the potentially widely varied processing objectives of an organisation.

The PRM server maintains several databases. A rights database provides information regarding how personal data is managed within the system. There are also databases containing data processor and data subject reference information, as well as activity logs for collecting information regarding the operation of the PRM system. Interestingly, while there is the potential for unbridled user tracking in a digital rights management system, in PRM, the tables are turned; the data processor and data controller activities are monitored. Any data subject tracking that could be added to a PRM system may only be in accordance with controlled application of Article 7 from the Directive.

When comparing Figure 1 and Figure 2, definite parallels may be drawn between the PRM and DRM system components. Table 2 delineates these comparisons.

Table 2. Comparison of DRM and PRM System Components

PRM Component	***PRM Comments***	***DRM Component***	***DRM Comments***
Data Subject	The Data Subject entrusts a relatively small amount of data for management by the data controller. The data controller manages the data, including its distribution to data processors.	Owner	The Owner entrusts its electronic property to the DRM server for distribution. In contrast to its PRM counterpart, the Data Subject, there are relatively few Owners as compared to Data Subjects.
Data controller Web Server	The data controller acts as the enforcer of usage requirements associated with personal data, with accountability provided through detailed logging. The web server provides an interface allowing data subjects several views such as the 'objection view' where they can access, rectify, revoke and maintain their personal data. Data controllers are provided with management views of the PRM system operation.	Content Provider Web Server	The Content Provider Web Server provides an interface allowing owners to maintain personal data and for management of the system. The Owners may track usage and other information regarding the their data.
PRM Server(s)	Privacy rules implementing the triad entity rights, preferences and requirements are handled here.	DRM Server	The DRM server contains rules implementing the way in which owners' property and subscribers' interactions are managed.
Personal Data	Data provided by the data subject that is traceable to them in some way.	Asset	The electronic property (content) entrusted to the DRM system for controlled distribution by the owner.
Protected Property	PRM protected property is personal information entrusted to the data controller, held and distributed using data protection. The number of entities among which the property is shared (data controller and data processors) is smaller than in the DRM protected property scenario.	Protected Property	Protected property is held and delivered using data protection. Access to the property is controlled via a rights usage policy. There may be a very large number of people gaining access to the protected property.

At the organisational level, there are also important distinctions between DRM & PRM. System elements within PRM models may well be operated by different legal entities. Thus the partner selection criteria for a privacy conscious firm will naturally consider the degree of trust a potential partner presents regarding its privacy practices. One foresees several ways to achieve those credentials. The most obvious way is extensive notification of organisational privacy practices, augmented with strong controllability via external privacy auditing from a reputable firm. For instance, one aspect of Dutch data protection interpretation of the security stipulations from the Directive is that contracts between data controller and data processor must provide assurance that the data processor will enforce a security policy that is at least as rigorous as the one to which the data controller is subject. Service Level Agreements that include bi-lateral audit rights are well applied here.

4.1 DRM evolution to PRM

The three aspects of DRM functionality of particular interest to PRM architectures are Asset Creation, Asset Management and Asset Usage.

Asset Creation (as illustrated in Figure 3) supports rights creation and validation. Rights validation ensures content may only be created from *existing* content if the rights exist to do so. Rights creation allows rights to be assigned to content. Below we illustrate asset creation functionality.

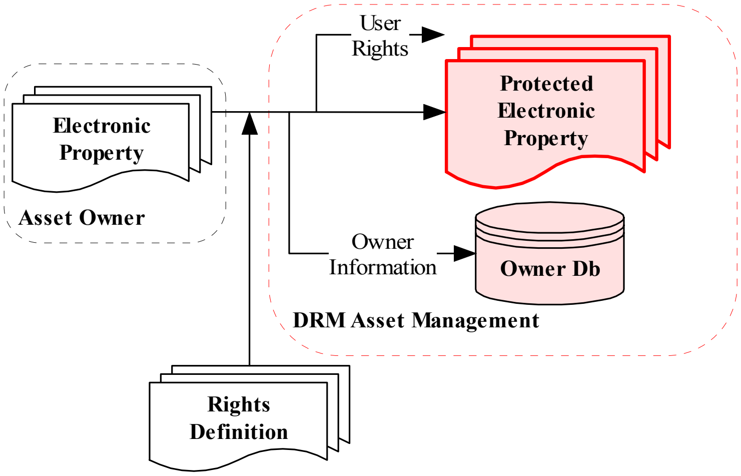

Figure 3. DRM Asset Creation

The driving purpose behind DRM: content distribution management, relates very well to data protection constructs constraining the onward passage of personal data. Article 6 (d) of the Directive build arguments related to the responsibility of data quality on the part of a data controller and data processor. Similarly, Article 12 (b) requires the data controller to provide the data subject with the opportunity to amend their personal data. In addition, the data must be of consistent quality throughout (for instance, when it is distributed among all data processors). The Directive also calls for a retention period of subject data that is either based upon legitimate grounds or consented to by the data subject.

Asset Management supports the access and retrieval of both content and metadata in distributed databases. Asset Management also provides logging functionality. Article 6 (c) requires data controllers to process volumes of personal data that are minimised for the task at hand. More centrally, PRM asset management permits data subject rights over their data (for instance for data versioning. However, a PRM asset management rules engine also codifies data controller interests so that, a data subject objection to a processing request may be rejected by the data controller, if the data subject had previously consented to the processing.

The PRM system must implement a high degree of monitoring of subject data usage. As well, the monitoring process itself must be protected. In most cases, there will be multiple data processors operating on personal data. Each of these data processors should maintain cryptographically protected log files [15] relating to the operations on every individual’s personal data. This approach would embed the accountability required of the data controller in the PRM system. In addition to meeting the requirements of the Directive, it would offer the data controller a means for monitoring of its privacy performance via log analysis.

Asset Usage supports permission management and (depending upon definition) audit trail functionality that would allow the usage environment to honour rights associated with content. This provides a means for monitoring and tracking content use. Below we illustrate some functional elements of a DRM system required for content usage monitoring.

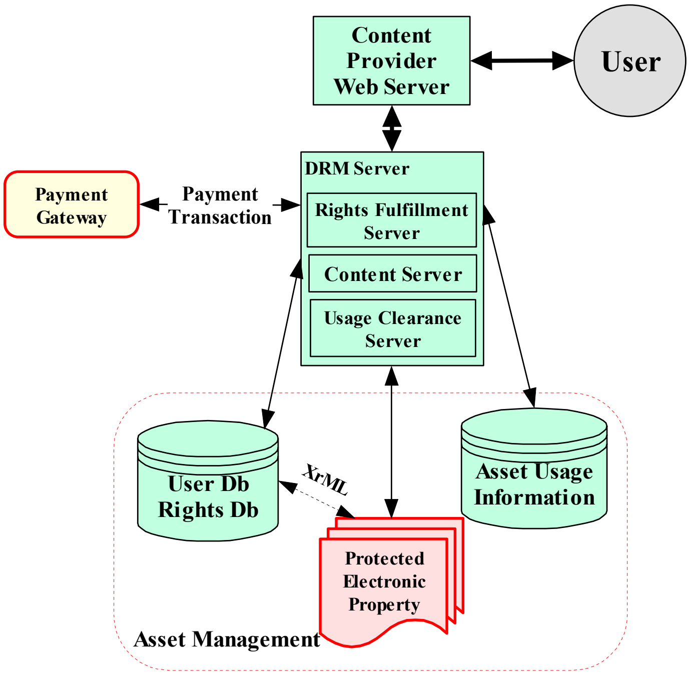

Figure 4. DRM Asset Usage.

The PRM server must extend the core logic associated with DRM server asset usage so as to support the PRM operational context: a large number of different owners of electronic property, many distributed data processors, and major responsibilities under the Directive. Three key entities contained within a DRM server are the Content Server (CS), the Rights Fulfilment Server (RFS) and the Usage Clearing Server (UCS). These entities are present, but reconfigured, in a PRM server. In DRM, the CS standard task is to distribute cryptographically packaged content, accessible by retrieving content and rights keys. In PRM, the situation is quite similar to DRM: the CS manages the controlled distribution of personal data assets. One difference however is that the secure container may have a varied granularity level for asset protection and auditing requirements based upon role-based rules dictating and auditing access on the grounds of consent or other Article 7 permission.

The functionality provided by the RFS in DRM ranges from providing payment receipts to recording asset accesses and devices used for those accesses. In PRM, the RFS enables the tracking of data processor use of subject data. The Asset tracking databases must be tamperproof, to prevent unauthorised changes to the tracking records.

Article 6 (b) of the Directive may be implemented by appending a retention period to personal data. This retention period is transfer-independent. If the period agreed to is 30 days, and the data controller passes this data to a data processor after 15 days, the data processor must conform to the remaining 15-day retention period. Once the retention period is exhausted, all instances of the personal data must be erased. Given this requirement, RFS functionality may be extended to coordinating asset usage information databases. This extension is required to meet Article 6 (b). To support temporal semantics, a secure timing mechanism linked to the database is required.

Basic UCS functions include recording and analysing transaction data. From Figure 4, the DRM server, when operating as a PRM server is advised by the rights database of the degree to which personal data may be disclosed to other parties, according to original data controller data capture conditions. Operation of the usage clearance server must link different purpose specifications to different parts of subject data. This permits implementation of both Article 6 (b) through the ability to identify which (element of) data is needed for each purpose, and Article 6 (d) via the retroactive and proactive updates necessary to assure data accuracy (plus audit trailing) in the relationship between data protection concepts such as purpose specification and the personal data itself.

Underlying these PRM requirements is a concept of data subjects controlling their personal data in much the same way as content owners or distributors control and monitor access and use of their digital assets with DRM. Interestingly, DRM systems gain maximum leverage from personal data through tracking consumer activity to any degree and subjecting that output to data mining at a clearing agency. For PRM, in comparison to DRM, the amount, quality and granularity of tracking information generated for digital media are extensive.

5. Expressing Privacy

In this section, we describe entity modelling for a PRM system. The Extensible Rights Mark Up Language (XrML) is a standard vocabulary for expressing the terms and conditions for the use of assets [16]. XrML Specification 2.0 proposes a base set of semantics useable for PRM purposes including rights holders and the expression of permissible usages for assets. Consider below a Data and Entity Model for XrML highlighting several PRM related entities.

XrML 2.0 is architected as a Core with context definable extensions using XML Schema. In Figure 5, the XrML entity model is presented as the Core and two Extensions (Standard and Content). Using this architectural approach, PRM may develop its own specific extensions, leveraging the work of the other extensions and a common Core.

DRM rights describe permissions, constraints and obligations between users and content. DRM systems rely on client software receiving rights, formatted in rights languages, expressing the limits or conditions for content use, e.g. the number of times content can be used. Rights metadata defines control over such content - for instance *view* but not *edit*. In PRM, the data subject may configure these rights so as to express control over her personal data, once it is under the auspices of the data controller.

XrML *Grant*, *Rights*, *Principal*, *Resource* (content) and *Condition* abstract elements - as they relate to the rights entity in the specification's data dictionary - in their current form all possess syntax that is of use to PRM. Additional data dictionary elements required for PRM may be developed through XrML Extensibility specifications. The *Grant* element represents a relationship of the entity's asset (XrML *Resource*), the entity itself (XrML *Principal*), permissions (XrML *Rights*) and context (XrML *Conditions*) so as to express agreements between data controllers and data processors for specific rights over personal data. Specification of defined expression containers and linking is one mechanism for generating data protection service level agreements (SLAs) between different legal entities operating in the PRM system.

A fundamental aspect of XrML, propagating PRM, is the notion of a "lifecycle" of rights, in that every distribution point of the data can be modelled. This "lifecycle" starts at the owner of the asset and follows the asset through multiple distribution possibilities to the final end user of the asset. Each distribution point must comply with the rights expression thereby assuring the trust integrity of the entire chain with respect to the wishes of the original owner. For example, a data subject may issue a *License* (an XrML term for a complete expression) to a data controller to issue *Licenses* to its asset. The *License* authorizes the data controller to issue certain rights to data processors if certain conditions are met. Such capacity offers transparency and control personal data use in a distributed environment.

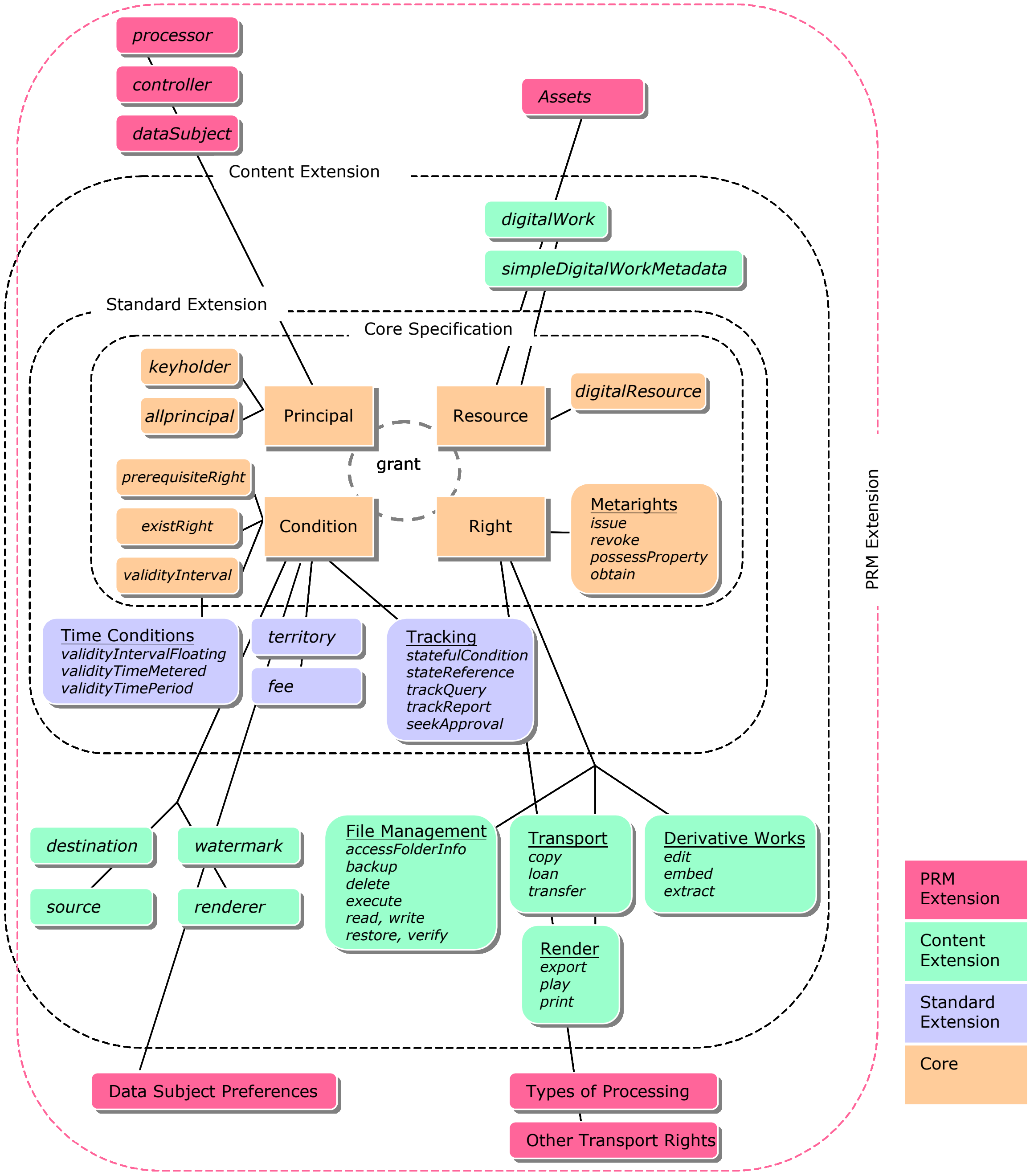

Figure 5: XrML Data and Entity Model with Sample PRM Related Entities, with the key PRM extensions highlighted.

Modelling content is necessary in PRM because personal data is a heterogeneous asset in terms of its sensitivity and the exercise of data subject control over (some) aspects of it. Descriptive metadata for the data subject indicates that standards are required for the degree of granularity required for both the content and the tracking information.

In terms of the *Principal* expression, PRM is primarily interested in the multiple data processors. All data processors must enforce and be advised of the processing preferences and requirements for assets, as denoted by data controller jurisdiction and as constrained by self-determination metadata constructs appended to those assets. The jurisdictional requirement regarding rights implies that in a PRM system a "policy evaluation" must be made before granting rights to a data processor. In other words, there is a "lifecycle" for the rights as "grantable" rights are transformed into "granted" rights. In essence, rights for any personal data must match the legal requirements of the country of origin for the data controller.

Regarding *Rights* expression (see Figure 6), PRM indicates any extension needs to consider a vocabulary describing such factors as: access rights to profiles, data subject's metadata profile in terms of granularity and tracking extensiveness, as well as the contingent responsibilities passed onto interacting data processors. Within the current *Rights* abstract element defined in the XrML Core and Content Extension, *File Management Rights*, *Render Rights*, *Derivative Works Rights* and *Transport Rights* elements have direct applicability for PRM.

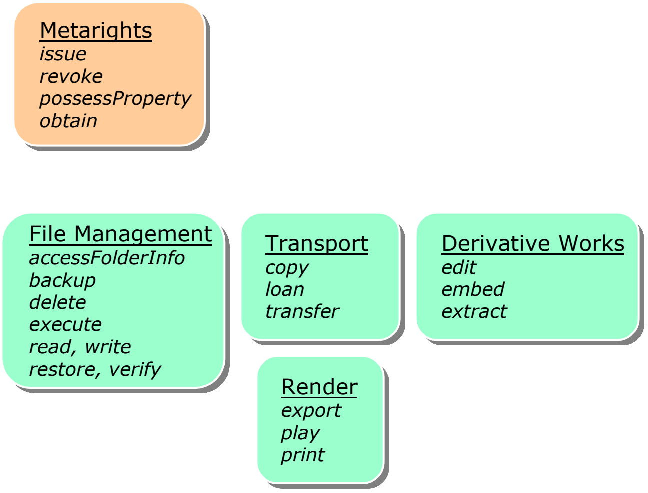

Figure 6: <Rights> Abstract Element and the elements defined in the Core and Content Extension

These rights definitions may be drawn upon in a way that gives the data subject an unprecedented level of control over the processing of her data by disparate data processors within a PRM system. The *render and derivative works rights* indicate syntax applicable to reuse of (some part of) personal data. The *edit and loan* abstractions would facilitate the communication of versioning responsibility upon the data controller and assist the enforcement of that responsibility upon every implicated data processor.

Metarights in the XrML Core offer the ability to model the "life cycle" of the rights. This may be applied for benefit of the data controller to model a controlled onward transfer of personal data. Indeed, the data subject herself may also make use of such functionality, hence offering the potential for a distributed strategy optimised in terms of privacy rights, preferences and service goals on the part of the data subject, and legitimate data controller interests.

Within the current *Conditions* abstract element defined in the XrML Core and Content Extension, several elements have particular applicability for PRM as shown in Figure 7.

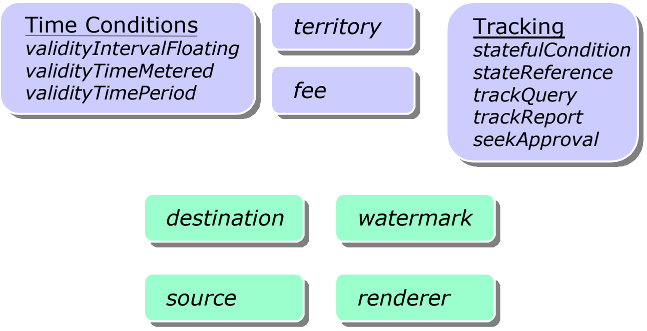

Figure 7: <Conditions> Abstract Element.

The *time conditions* element is a useful definition for PRM, so long as different time granularities (e.g. seconds, minutes, hours, days, months, years) are supported. The retention period functionality discussed earlier implies the need for a timestamp. The *territory condition* is an important tag to designate the country of origin of the data controller in PRM. With an indication of EU location, personal data, irrespective of the data subject's nationality, must conform to the control and processing restrictions related to the particular national law of the data controller. If the indication is of US location, wherein only US data subjects will be dealt with, there is currently no legal requirement to implement PRM system functionality, unless the processing intentionally makes use of equipment based in the European Union under Article 4-1(c).

The *destination condition* is interesting because it can provide the semantic basis for the aspect of Article 6(b), which limits the transfer of assets, even within the same legal entity, to uses that are similar to the original purpose of processing. Using various expressions, one can develop a definition of quality that is customized to the context. For example, the *watermark condition* can be specified such that a printout is required to have a watermark indicating its origin or identification. The *renderer condition* may also be used to specify a particular application that can faithfully render the information. Versioning may also be expressed through the use of the *seekApproval condition*. Finally, the *tracking condition* can also be used throughout to enable other types of checks.

6. Discussion

To present an analysis of the potential DRM technology as a tool for privacy protection, we first analyse PRM requirements and implementation challenges with respect to the facilitation principles. Following this analysis are some additional comments and descriptions of challenging research issues in the further development of PRM.

Principle 1: Reporting the Processing

The PRM server block tracks the data processors with which the data controller web server has processing arrangements. Of interest are who does the processing, the type of processing, and on what assets processing is being applied. Adaptation of DRM asset tracking functionality can achieve these requirements.

There are many possible data subjects (many millions) and data processors (many hundreds both local and remote). This contrasts highly with DRM wherein there are a limited number of content owners (Data Subject counterpart) and many millions of subscribers (data processor counterpart), only a limited number of whom would be active at a time. In the case of PRM, all of the Data Subjects may expect reports, and there may be many hundreds of data processors active at any time. Therefore despite individual asset size being small, highly scalable databases are required to manage reporting processes. Scalability of the PRM server architecture must be addressed as well to deal with the enormous numbers of simultaneous logging and reporting activities.

Principle 2.: Transparent Processing

The PRM server provides data subjects the ability to view data controller / data processor operations on subject data on an "on demand" basis. DRM systems currently meter usage of content. In a PRM system, adaptation of usage tracking through secure distributed logging techniques is required. Centralized management via asset usage monitoring is one approach, although this may not scale well. A major challenge is that it is not currently possible to determine exactly what a data processor is doing or has done with the personal data ir has received.

Principle 3: Finality

The DRM notion of asset rights offers a means for specifying and enforcing requirements on processing of personal data such as data subject consented retention periods. There are many data subjects, with a minimal level of commonality in rights specification and a degree of uniqueness to their separate rights specifications for processing. Since the data controller holds the rights, and there are many distributed data processors that must access each data subject's rights for processing, subject data must be checked out from the data controller whenever it is processed. This presents a scalability challenge. One way to mitigate this challenge would be to distribute rights with personal data. Functionally (if not legally), this distributes responsibility for data protection enforcement from the PRM server to the data processor. A means for maintaining data controller linkable responsibility is facilitated by the rights granting model specified by XrML. Personal data can travel separately with only information on where to get permission to process. At the time of processing, a request can be made to the data controller, which would in turn return an XrML "License". The "License" would contain the permissions and the conditions (time, territory, tracking state, etc) under which the permissions can be consumed.

Principle 4. Lawful basis for data processing

The data controller may only process personal data on certain grounds, grounds, which must be replicated in all data processors of the personal data. A central problem of data processor processing enforcement has yet to be solved. The first step, however, is to standardize a data protection definition language as a starting point so as to control parsing. However, XrML provides a means for specifying credentials that the data processor must possess in order to grant permissions. For example, XrML may specify that the data processor possess a digital certificate signed by some authority. The system then validates the credentials of the data processor prior to sending a license that can be used to decrypt and the protected data. Ultimately, the effectiveness of this approach depends upon the trust between the data controller and data processor, since once the data is in the clear, an unscrupulous data processor may use it for any purpose. To a certain extent, watermarking techniques may be used to mark private information in order to determine the origin and originality of the data. However, watermarking approaches are generally considered to offer weak security [17]

Principle 5. Data Quality

Quality relates to specified attributes the asset must maintain. Effectively it must be as correct and as accurate as possible for all who deal with the data. If the data controller maintains a central repository of subject data and controls access for each data processor request, there is a reasonable likelihood data quality may be maintained. However, this approach is not scalable. If the personal data is distributed to provide scalability, the data must first of all be protected, and secondly, it must be possible to assure the data is consistent throughout.

Principle 6: Rights

The data subject has the right to determine and maintain the correctness of the relevant personal data held by the data controller. The data subject also has the right to raise objections as to the behaviour of the data controller and data processors. This requires editing provisions and a communication channel for raising objections.

In DRM asset management, owners may transfer content to the server for distribution, but no on-line tools exist currently for editing content – such tools must to be added to the PRM server to support an editing function. Access to the editing function must be authorized. As well, a communication channel for raising objections is required. An effective tool for raising objections would include data mining tools of processing transactions. The objective is to provide evidence of contract, or privacy breaches.

Principle 7: Data Traffic Outside the EU

The PRM server block enforces grounds for data transfer on the basis data adequacy and exceptions – for EU nationals of different member states, as well as say American nationals who express a self determination for EU data protection applied by a PRM system.

This requires the ability to identify the nationality of the data controller, and data processors, and the enforcement of suitable logic appropriate to origin. There are currently no foolproof technologies to determine geographic location of users. As well, rules to support multiple countries would be extraordinarily complicated.

Principle 8. Data processor processing

The PRM server must decide when it will outsource data processing, on the correct grounds in a dynamic arrangement. Enforcement of data controller rules is vital. It is clearly challenging for the data controller to enforce processing among widely distributed data processors, apart from recourse to third party auditing. Negotiation techniques between data controller and data processor could determine a likelihood of compliance, but not enforcement.

Principle 9: Security

The data controller is responsible for ensuring data processors apply uniform security standards across all data controllers. DRM secures content for distribution – PRM builds on this in an adaptive way as data protection prescribes – such as relating authentication to data sensitivity.

General Discussion

While it may appear that some principles are easily accommodated without considerable development to extend DRM systems, there are areas that require further research and development. For instance, protection against unlawful processing and data traffic outside of the EU are two key areas where new technologies must be developed. In the case of the former, a means for tracking the actual processing performed by a data processor is needed. A DRM system is well suited to track the time a data processor requests and receives data for processing, however it is not designed to track and record the actual processing performed. Once a data processor receives the data in the clear for an expressed purpose, the data processor may simply do what it wishes with data. This information "leakage" by data controllers or data processors would be difficult to detect. To remedy this situation at least two approaches may be taken: develop a means for determining the actual processing done by the data processor or implement and use a reputation management and reporting system to determine which data processors may be most trusted to deal with private data. One possibility for accomplishing processing management might be a specialized type of sandbox for subject data processing wherein the data subject data would be entrusted to the data processor only if the processing to be performed by the subject may be verified before and after processing operations.

With respect to the issue of data traffic outside the EU, one aspect of this issue is the ability to determine geographical location of data controller or data processor representations. There have been techniques and at least one service [18] developed to determine the geographical location. Unfortunately, these approaches are far from foolproof. One means of circumvention involves the deployment of dynamic proxies. A further complication to dealing with data traffic outside the EU is that privacy laws do not have consistent electronic implementations that would facilitate any sort of automatic negotiation or decision-making around how to deal with subject data.

Other challenges exist regarding DRM extensibility into PRM: third party tracking, scalability; and the privacy issues associated with DRM systems. Regarding third party tracking and scalability, DRM was developed to support delivery and protection of potentially vast amounts of electronic property from typically just a few owners or distributors. In PRM, relatively small amounts of data are collected from a very large number of citizens, where the citizen entrusts information to a data controller. Citizen data must be tracked for use. This data be managed, kept confidential and be editable by the citizen to assure accuracy. A PRM system is designed to keep the data protected as well as track the sharing of personal data. It is clear that conventional DRM systems will require extensive redesign to support this demanding application. Incorporating a Trusted Third Party approach wherein, a data controller or data processor must "check out" information each time it is used may appear to offer a solution to this issue. Unfortunately, this approach adds considerable overhead to data controller and data processor activities as well as being a single point of failure. An alternate approach might be the delegation of the use-tracking function to the data controller. While this would distribute the tracking function load, it would also require a high degree of *trust* between the owner and the data controller.

DRM systems are not without controversy regarding privacy. Since DRM systems track what users purchase, how often they access material, when they use it, it is clear that these systems may be used to track detailed activity of subscribers [19]. This makes it possible to substantially profile individuals. With this potential for privacy breach one may wonder the value of considering such systems to implement privacy rights management to uphold data subjects' privacy. In the PRM we describe, the tables are turned; the digital material of value is user data. It is treated similarly to the digital assets under management in DRM with PRM tracking the use of personal data by data controllers.

Another issue concerning adaptation of DRM systems for PRM is an analysis of system adversaries. For DRM, the users form the adversaries. Skilled users may (and have) found ways to subvert systems that protect delivery and/or playback of electronic property such as media files and e-books. The approaches range from taking advantage of vulnerabilities in software or operating systems to accessing drivers, for instance, audio drivers, to divert digital audio streams to storage, in the "clear". In some cases, easy-to-use, software has been made available on the Internet allowing even relatively computer illiterate users to gain access to unprotected digital media.

While this is a particular exposure in DRM, for Privacy Rights Management, the issue is quite different. The protected material is the citizen's private information. The data controller protects this private information. It is only exchanged with the citizen when it is entered or updated. Other individuals have little interest in the data. Data controllers and/or data processors however may wish to process and sell information from large numbers of citizens for purposes other than those that were deemed appropriate by the citizen and the data controller. It would be difficult to prevent this with today's open computing architectures. There are however several factors that may regulate data processing.

First of all, the reputation of the data controller and data processor are at stake. Currently, public disclosure of any leaks of private information, from web sites for instance, results in loss of reputation to the company owning the web site. Generally, business is affected by such a loss of reputation. Data controllers and data processors that are serious about staying in their respective businesses will do everything possible to maintain reputation. In addition, depending upon jurisdiction, there are significant deterrents to misappropriation and misuse of private data. Large fines may be levied against offenders.

There are also some developments that might counter this type of attack. In mid 2002, Microsoft announced Palladium [21], a type of trusted computing platform [22]. While it is difficult to determine the exact details of the system Microsoft is planning, some of the features supporting digital rights management could be quite useful for privacy rights management. A truly secure computing platform might reduce the threat private data exposure to unscrupulous data processors. A secure computing platform, for instance, could make enforcement of the legal use of data possible. For instance, a data controller might require data processors to use a secure computing platform to access personal data. This would make extracting the data for other uses more difficult.

As we have illustrated, simply protecting data in storage and transit is no longer enough when considering the scope of privacy laws like the Directive. We propose an augmented abstraction of DRM functionality to provide privacy rights management functionality. Given the embryonic commercial status of the privacy market, a PRM investment appears worthwhile both in terms of what is necessary to come close to achieving compliance with current and emerging legislative requirements, and what is required to meet corporate privacy policies towards building a stronger trust relationship with clients. On the other hand, while the application of digital rights management appears to offer promise for privacy rights management, a full implementation to support the Directive would be challenging. While this is a preliminary report on this new approach for privacy enforcement, our future work will focus on the development of specific technologies to address the issues of reputation management, scalability and data processor monitoring and process limit enforcement.

Acknowledgements

We would like to thank Alexander Dix, Brandenburg Privacy Commissioner Germany, Anne-Marije Fontein at the Dutch Data Protection Authority, Adam Stostack of ZeroKnowledge Systems, Mike Gurski, Information & Privacy Commissioner, Ontario, Helena Lindskog at Ericsson. Also we have very special thanks to Hari Reddy (OASIS Rights Language Technical Committee, Chairperson) and Guillermo Lao of ContentGuard for their help with XrML extensions for privacy rights management.

References

[1] Official Journal L 281, 23/11/1995 p. 0031 – 0050.

[2] B. Rosenblatt, B. Trippe, S. Mooney, Digital Rights Management: Business and Technology, John Wiley & Sons, ISBN: 0764548891, 1st edition , November 15, 2001

[3] eMeta Corporate Web site at: http://www.emeta.com/

[4] Digital Owl Corporate Web site at: http://www.digitalowl.com/

[4] Dot Encrypt Corporate Web site at: http://www.dotencrypt.com/

[6] Alchemedia Corporate Web site at: http://www.alchemedia.com/

[7] Authentica Corporate Web site at: http://www.authentica.com/

[8] Windows Media Rights Management at: http://msdn.microsoft.com/library/en-us/wmrm/htm/windowsmediarightsmanagersdk7.asp

[9] IBM EMMS Web site at: http://www-4.ibm.com/software/is/emms/

[10] Intertrust Corporate Web site at: http://www.intertrust.com/

[11] Perimele Corporate Web site at: http://www.perimele.com/

[12] Viaquo Corporate Web site at: http://www.viaquo.com/

[13] Council of Europe Convention 108: Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data. Available at: http://conventions.coe.int/treaty/EN/Treaties/Html/108.htm

[14] Stefan Brands, *Rethinking Public Key Infrastructures and Digital Certificates; Building in Privacy*, August 2000, MIT Press

[15] See B. Schneier, J. Kelsey, *Secure audit logs to support computer forensics*, ACM Trans. on Information and System Security, Vol. 2, No. 2, 1999, pp. 159-176.

[16] XrML is being contributed to the standards body OASIS Rights Language Technical Committee as its foundation technology. More information can be found at http://www.oasis-open.org/committees/rights/ or at http://www.xrml.org

[17] I.J. Cox, M.L. Miller, Electronic Watermarking: the first 50 years, 4th IEEE workshop on Multimedia Signal Processing, Cannes, France, Oct. 3-5, 2001, pp. 225-230.

[18] Quova, Inc. at: http://www.quova.com/

[19] J. Feigenbaum, M. Freedman, T. Sander, A. Shostack, Privacy Engineering for Digital Rights Management Systems, Oric, of the ACM Workshop on Security and Privacy in Digital Rights Management 2001. Available at: http://citeseer.nj.nec.com/feigenbaum01privacy.html

[20] Open Digital Rights Management, at: http://www.odrl.net/

[21] A. Carroll, M. Juarez, J. Polk, T. Leninger, Microsoft "Palladium": A Business Overview, at: http://microsoft.com/PressPass/features/2002/jul02/0724palladiumwp.asp

[22] Trusted Computing Alliance at: http://www.trustedcomputing.org/tcpaasp4/index.asp

[23] S. Garfinkel, D. Russell, Database Nation: The Death of Privacy in the 21st Century, O'Reilly & Associates, ISBN: 0596001053, 2001, 336 pages.