LongRec

Compliance Work Package

State-of-the-Art

Hannelore Dekeyser

September 2008

ICRI - K.U.Leuven - IBBT

This report is produced as a contribution to the LongRec (Long-Term Records Management) project headed by Det Norske Veritas (DNV) in collaboration with a number of case partners, commercialization partners and research partners. The primary objective of LongRec is persistent, reliable and trustworthy long-term archival of digital information records with emphasis on availability and use of the information. The project's public web site is at `http://www.longrec.com`.

LongRec is a three year project (2007-2009) partly funded by the Norwegian Research Council. The project constitutes the Norwegian team of the InterPARES 3 project, `http://www.interpares.org`

Contents

Chapter 1

Compliance

Compliance is "conformity in fulfilling official requirements"[1] or "demonstrating conformity with regulatory or legal constraints".[2]

Achieving compliance is challenging for any organisation, and even more so for multinationals. Firstly, an overview of all relevant regulatory provisions must be obtained and kept up to date. In particular, sector specific regulations must be taken into consideration. Secondly, the – usually abstract – regulatory provisions must be interpreted and applied to the specific situation of the organisation. Interpretations may evolve over time, notably influenced by administrative decisions or jurisprudence, making it necessary to keep track of which interpretation whas held at what time. To complicate matters further, the entire body of regulations in a country or region is not necessarily coherent. Only too often do gaps or even contradictions become apparent when applying the law. Multinationals are per definition subject to rules stemming from different jurisdictions, the odds of such an organisation not running into conflicting regulations is close to zero. Ultimately, no matter how complicated the set of legal rules applicable, the organisation must draw its conclusions in order for it to effectively design its business processes and set out policies that determine its course of action.

The process of preparing compliance policies as well as monitoring and enforcing them is not automated to a significant degree.[3] Usually, inhouse legal advisors and/or external legal counsel

[1] Online Merriam-Webster Dictionary, `http://www.m-w.com`.

[2] ROWLINGSON, ROBERT, "A Ten Step Process for Forensic Readiness", *International Journal of Digital Evidence*, vol. 2 2004, Nr. 3 , `http://www.utica.edu/academic/institutes/ecii/publications/articles/A0B13342-B4E0-1F6A-156F501C49CF5`, p. 10.

[3] There are of course companies that advertise their solutions as 'ompliance' solutions. It is difficult to tell whether this is more than another way to sell their IT security solutions to companies. Two (random) examples are `http://www.agiliance.com/` and `http://www.safestone.com/`.

seek out relevant regulations as well as their interpretation as input for designing business processes and drafting compliance policies. At a later time, either on a regular basis or as the need presents itself, the business processes and compliance policies may be reevaluated and modified in light of modifications to regulation or changed interpretation of regulation or in light of evolving business needs.

A great many countries systematically publish legislation and regulation online and have done so for quite some time.[4]Increasingly, countries not only offer a chronological view of issued regulation, but also make available consolidated views on regulatory texts. The consolidated view is often limited to the version in force at present, but more and more systems are capable of showing point-in-time consolidated versions.

An extensive overview of online sources of public legal information[5] is maintained by the World Legal Information Institute (WorldLII).

Such regulatory databases are a great tool for human users, but a far cry from machine-readable texts. Current research in the field of legal informatics and information retrieval is devoting much effort to making machine-readable texts a reality.[6]

Which records management system to use for policy documents is a matter each organisation must decide for itself. From a compliance viewpoint, backlinks from policy to the regulatory texts they implement are of great importance. From an evidence viewpoint, the preservation of policy documents over time in an authentic way is imperative.[7]

Having compliance policies is one thing, their application in practice is quite another. A first step to take is to ensure that the organisation's records, stemming from its – supposedly compliant – activities, should appropriately reflect which policies were relevant to them and ideally also show that they were followed. Linking records with policies, as well as recording data about their application, is a matter for legal metadata.[8] To illustrate the need for legal metadata, consider records containing personal data[9] and/or copyrighted information.[10] All records con-

[4] It should be noted that in many jurisdictions only the official publication on paper is the authoritative version of regulatory texts. In case of a discrepancy between the paper and online version, the paper version takes precedence.

[5] "Public legal information means legal information produced by public bodies that have a duty to produce law and make it public. It includes primary sources of law, such as legislation, case law and treaties, as well as various secondary (interpretative) public sources, such as reports on preparatory work and law reform, and resulting from boards of inquiry. It also includes legal documents created as a result of public funding.", Declaration on Free Access to Law, `http://www.worldlii.org/worldlii/declaration/montreal_en.html`.

[6] See 2 "Legal Informatics" on page 6.

[7] See 3 "Evidence" on page 25.

[8] See 4 "Legal Metadata" on page 42.

[9] See 4.4.1 "Privacy Metadata" on page 43.

[10] See 4.4.2 "Copyright Metadata" on page 60.

taining personal data should be linked to their governing privacy policy or policies, which detail how they ought to be handled. If the privacy policy demands that consent is obtained, it is imperative that metdata reflects that such consent was in fact obtained. If an exception is invoked in order to process data without consent, this should be recorded in stead. The same holds for copyrighted records. A link to the copyright policy serves to show how the records ought to be treated throughout their life-cycle in an organisation. Metadata should record how the policy was applied in practice, e.g. obtaining a licence, operating under a copyright exemption, etc. Ultimately, it serves little purpose to link records to policies if there is no way of knowing whether or not the policies were actually followed in any given instance.

Compliance is not so much a separate activity, as a quality of the way in which an organisation conducts its activities. The design of work processes should lead to compliance, and the records resulting from those processes should reflect their compliance. In turn the records should be handled in a way that is compliant with rules applicable to records management as a work process in its own right.[11]

How can information technology improve efficiency in the process from regulation to policy formulation and ultimately compliance?

Two quick wins can be identified: deploying an information management system for regulatory texts and implementing a records management solution for policy documents. Ideally, both these systems

should be able to give point-in-time consolidated versions of either regulations or policies.

Creating stand-alone systems for regulations and policies respectively is only a small step forward. The next step is to interconnect these knowledge bases in the most efficient way. A relatively straightforward way of doing this is creating backlinks from policy documents to the regulatory texts from which they stem or of which they are an application. Preferably the links between policies and regulations is very precise, on the provision or even rule level.

A more sophisticated form of backlinking would take into consideration intermediate steps which inform the organisation's decision making process from regulation to policy, notably jurisprudence, legal doctrine, and the opinion of inhouse or external legal counsel.

Aside from further research in the field of legal informatics, there is an economical aspect to consider. Developing and maintaining such sophisticated legal information systems as are envisaged is very costly. Even large multinational corporations may find the costs prohibitive. A business modeling exercise could examine alternative approaches, e.g. inhouse development,

[11] See 5 "Recordkeeping" on page 93.

outsourcing, community platform building, etc. Figuring out which model(s) are economically viable would allow interested organisations to choose a winning strategy.

Further reading

- Governance & Compliance Project `http://www.zurich.ibm.com/csc/security/compliance.html`, IBM Research Laboratory
- GASSER, URS and HAEUSERMANN, DANIEL M., "E-Compliance: Konzept, Merkmale, Aufgaben und organisatorische Auswirkungen", In X. (ed.), *Internet-Recht und Electronic Commerce Law: 9. Tagungsband*, Bern, Stämpfli, 2006
- PFITZMANN, BIRGIT, POWERS, CALVIN and WAIDNER, MICHAEL, *IBM's Unified Governance Framework (UGF) Initiative*, Zurich, IBM Research Division, 2007, IBM Research Report RZ 3699 , `http://www.zurich.ibm.com/pdf/csc/rz3699_UGF-whitepaper.pdf`
- GASSER, URS and HAEUSERMANN, DANIEL M., *E-Compliance: Towards a Roadmap for Effective Risk Management*, Harvard, The Berkman Center for Internet Society, 2007 , `http://ssrn.com/abstract=971848`

Chapter 2

Legal informatics

Legal informatics studies the application of information technology to the practice of the law:

> "[L]egal informatics is a complex of theories, methods and techniques for getting to know, producing, managing, amending, using law, with the aim of guaranteeing the certainty of the law itself, increasing the efficacy of the services offered by the justice establishment, contributing to making of law a special social function promoting the integration, the balance and the cultural evolution of the social system."[1]

> "Legal informatics is, then, that science concerned with problems linked to the effective storage, retrieval and transmission of legal data; but it also deals, and from a slightly different perspective, with problems relating to the rationalization of legal activity; within this second grouping, the studies relating to formalization of the legal order (in particular, research in the fields of legal language, formal legal logic and artificial intelligence in the field of law) take on particular importance."[2]

An overview of the developments in legal informatics so far, as well as current and future trends in research is provided in the One-Lex mission statement.[3]

[1] LIMONE, D. A., "L'insegnamento dell'informatica giuridica in Italia", In FROSINI, V and LIMONE, D. A. (ed.), *L'insegnamente dell'informatica giuridica*, Naples, 1990. English translation from BINAZZI, SIMONA et al., "ITLaw: An Advanced Documentation System in Legal Informatics", *The Journal of Information, Law and Technology*, 1999, Nr. 1 , `http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/1999_1/idg/binazzi/`.

[2] BINAZZI et al., *The Journal of Information, Law and Technology 1999*, *op. cit.* (as in n. **??**).

[3] `http://www.one-lex.eu/The%20Project/Mission/Mission.html`.

2.1 Access to law

The World Legal Information Institute (WorldLII) maintains an extensive overview of online sources of public legal information, meaning "legal information produced by public bodies that have a duty to produce law and make it public. It includes primary sources of law, such as legislation, case law and treaties, as well as various secondary (interpretative) public sources, such as reports on preparatory work and law reform, and resulting from boards of inquiry. It also includes legal documents created as a result of public funding."[4]

A basic problem to solve is how to identify norms and the individual provisions within them, as well as how to persistantly be able to point to their location.

In the political and economical landscape of today, access to the law must necessarily include access to legal texts stemming from the transnational and/or international level.[5] This implies finding an appropriate method to deal with the use of multiple languages.

2.1.1 Referencing legal norms

Unambiguous pointers are a prerequisite for a reliable system of cross-referencing between regulatory texts, both within one jurisdiction and amongst different jurisdictions. Such pointers would be equally usefull to link back to regulation from policy documents intended to implement these, which is a functionality of the compliance solution proposed here.

Research has sought to solve the referencing problem with the development of legislative XML schemas.[6] The challenges faced by any standard for legislative XML, in particular when several jurisdictions are involved, is discussed by Hatter.

Particular to the legal world is the importance of maintaining an overview of what happened over time. The law is not static but changes over time. Being able to reconstruct the current, past and future versions of legal texts is crucial in understanding which obligations and rights pertained at a specific point in time. The most consistent way to incorporate these transformations in metadata schemas is by using event descriptions, according to Boer, however he notes

[4] Declaration on Free Access to Law, `http://www.worldlii.org/worldlii/declaration/montreal_en.html` .

[5] HATTER, CLYDE, "Standard Models for Legislation - The Cost of Compliance", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art16.pdf`.

[6] An introduction to legislative XML can be found at the One-Lex project website: Presentations `http://www.one-lex.eu/Activities/summerschool/slides.html`, Papers `http://www.one-lex.eu/Activities/summerschool/readings.html`

that few metadata schemes for legislation include events as such.[7] A great number of dates may be of importance to a legal norm: adoption, publication, entry into force, efficacy, applicability, ... Sometimes laws come into effect retroactively or modifications only become or remain effective when certain conditions are fulfilled.[8]

Norme in Rete

The 'Norme in Rete' (NIR) project [Legislation on the Net] aims to create a single point of access to legal texts and to devise a mechanism of stable cross-references among legal documents. To achieve these goals, two standards have been developped:

1. a standard for cross-referencing legal documents using a uniform name (URN), which is an unambiguous identifier, allowing the references to be expressed in a persistent way, independently of document physical location;
   - The URN must be able to unambiguously identify any normative measure regardless which normative body issued it.
   - The URN must make the distinction between past, present and future measures.
2. a standard for legal document description by defining XML-DTDs (NIR-DTDs) of increasing degree of complexity.

The NIR URN syntax contains 5 elements:

- Name-space identified by 'nir';
- Enacting authority: eg Ministry of Finance;
- Type of measure: e.g. law, decree;
- Details: eg. Date of issue, Different later versions of the document;
- Annexes;

[7] BOER, ALEXANDER, "Using event descriptions for metadata about legal documents", In WINKELS, RADBOUD and FRANCESCONI, ENRICO (ed.), *Electronic Proceedings of the Workshop on Standards for Legislative XML*, 2007 , http://www.leibnizcenter.org/~winkels/events.pdf.

[8] PALMIRANI, MONICA and BRIGHI, RAFFAELLA, "Time Model for Managing the Dynamic of Normative System", In WIMMER, MARIA A. et al. (ed.), *Electronic Government*,; RIVERET, RÉGIS, PALMIRANI, MONICA and ROTOLO, ANTONINO, "Legal Consolidation formalised in Defeasible Logic and based on Agents", In BIAGIOLI, CARLO, FRANCESCONI ENRICO SARTOR GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art9.pdf; GRANDI, FABIO, MANDREOLI, FEDERICA and TIBERIO, PAOLO, "Temporal modelling and management of normative documents in XML format", *Data Knowledge Engineering*, vol. 54 2005, Nr. 3 , http://dx.doi.org/10.1016/j.datak.2004.11.002.

It should be noted that the URN system should be able to cope with changes in the state's organisation (new normative bodies, renaming of normative bodies, mergers and splits, ...). The NIR project uses a normalization mechanism to deal with variations or mistakes in the URN terms.

This project deals only with the legal texts of one country. This system could easily be adapted to an international context by introducing more name-spaces. For instance, country codes could be used for national legislation and specific codes for international and regional bodies (UN, OECD, EU, ...).[9]

In addition to the referencing system which allows to reliable locate legal texts, the NIR project has developed a set of XML schemas to structure legal texts internally.

Three DTDs of increasing degree of depth have been developed. The first is a flexible schema – 'DTD flessibile' (nirloose.dtd) – which has no mandatory rules so that it can be used for legacy legal documents. A second schema – 'DTD completo' (nirstrict.dtd) – is intended to support the drafting of new legal texts in accordance with certain drafting rules. A third schema – the 'DTD base' (nirlight.dtd) – is a subset of the 'DTD completo'.

The NIR-DTDs capture 2 different aspects of legal texts. Firstly, the formal characteristics of the text are described, there are elements for heading, preamble, sections, articles, paragraphs, references to other laws, tables, lists, ... This includes general metadata as subject classification, publication date and relationship with other acts. Secondly, the functional characteristics of the text are described. This is done by dividing the texts into its most basic components, being 'provisions' (a fragment of a regulation). The type of the provision can be described, most commonly 'obligation', 'prohibition', 'sanction', 'exception'. Particular metadata of an analytical nature is then added to these provisions, so-called 'arguments', which comprise information about the addressee, the counter-party, the action to be taken, ... Metadata regarding version control can be added as well, such as information regarding insertion of provisions, abrogation or substition.

Work is ongoing to incorporate time references into the DTD's. This relates to any reference to time within the text, eg. deadlines, prescription periods, etc., which will be described as an argument of a provision, but also to time constraints external to the text, e.g. entry into force or time of abrogation.

[9] See the solution proposed in SPINOSA, PIERLUIGI, "Expansion and Internationalization of the Italian Schema of Assignment of Uniform Names", *I Quaderni*, 2005, Nr. 18 , `http://www.cnipa.gov.it/site/_files/Quaderno\%2018.pdf` and SPINOSA, PIERLUIGI, "Internationalization of the Legal URN Schema", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art6.pdf`.

The NIR project started in 1999 and ended in November 2001, the results were published in *Informatica e Diritto*, 2000, vol. 1 and 2001, vol. 2. An XML editor using the NIR DTDs is under development.[10]

Further reading

- NIR project page `http://www.ittig.cnr.it/Ricerca/UnitaEng.php?Id=40\&T=1`
- BIAGIOLI, C. et al., "The NIR Project: standards and tools for the Italian legislative environment", Berlin, 2004 , `http://www.jurix.nl/index.php?option=com_docman\task=docclick\Itemid=27\bid=14\limitstart=0\limit=10`
- SPINOSA, PIERLUIGI, "Expansion and Internationalization of the Italian Schema of Assignment of Uniform Names", *I Quaderni*, 2005, Nr. 18 , `http://www.cnipa.gov.it/site/_files/Quaderno\%2018.pdf`, *op. cit.* (as in n. **??**)
- FRANCESCONI, ENRICO, "The "Norme in Rete"- project: Standards and tools for Italian legislation", *International Journal of Legal Information*, Vol. 34 2006, Nr. 2 , `http://www.xmleges.org/ita/images/stories/francesconiijli06.pdf`
- AGNOLONI, TOMMASO, FRANCESCONI, ENRICO and SPINOSA, PIERLUIGI, "xmLegesEditor: an OpenSource Visual XML Editor for supporting Legal National Standards", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art17.pdf`, *op. cit.* (as in n. **??**)
- SPINOSA, PIERLUIGI, "Internationalization of the Legal URN Schema", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art6.pdf`, *op. cit.* (as in n. **??**)

MetaLex

The objective of MetaLex is to create an open XML interchange format for legal and legislative resources.

[10] AGNOLONI, TOMMASO, FRANCESCONI, ENRICO and SPINOSA, PIERLUIGI, "xmLegesEditor: an OpenSource Visual XML Editor for supporting Legal National Standards", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art17.pdf`. See also `http://www.ittig.cnr.it/Ricerca/UnitaEng.php?Id=8\&T=3` and `http://www.xmleges.org/ita/`.

> "The standard intends to provide a generic and easily extensible framework for the XML encoding of the structure and contents of legal and paralegal documents. This obviously includes legislation and case law, but also written public decisions, internal and external business regulations (for instance ship classification rules as in [Winkels[11]]), and contracts. XML elements and structure are defined in schemas that can be used to validate a document. Since there is a great variety of legal documents that cannot be covered by one normative standard, the standard consists of multiple schemas defining vocabularies that can be mixed in a document."[12]

The following legal texts have been used as test cases:

- Dutch law on income tax in the context of the E-POWER[13]
- Dutch penal code of 1881 in the context of the e-COURT project
- Italian & Polish court room transcripts and case law for e-COURT

> "The standard differs from other existing metadata schemes for legal documents in two respects; it is language-independent and it aims to accommodate uses of XML beyond search and presentation services."[14]
>
> "The MetaLex XML schema aims to be a standard interchange format for legal documents for the purposes of presentation, description of the relations between legislative documents, search and filtering on meaningful levels of detail ([Moens[15], Turtle[16]]), and version management and file exchange."[17]
>
> "The MetaLex XML schema has been designed with multilingual regulations and differences between the main European languages in mind."[18]

[11] WINKELS, R.G.F. et al., "Generating Exception Structures for Legal Information Serving", In GORDON, TH.F. (ed.), *Proceedings of the Seventh International Conference on Artificial Intelligence and Law (ICAIL-99)*, New York, ACM, 1999

[12] BOER, ALEXANDER, HOEKSTRA, RINKE and WINKELS, RADBOUD, "Metalex: Legislation in XML", In BENCH-CAPON, TREVOR, DASKALOPULU, ASPASSIA and WINKELS, RADBOUD (ed.), *Legal Knowledge and Information Systems: JURIX 2002*, IOS Press, 2002 , `http://www.jurix.nl/pdf/j02-01.pdf`, p. 1.

[13] VAN ENGERS, T. et al., "POWER: Using UML/OCL for Modeling Legislation -an application report", In X. (ed.), *Proceedings of the 8th International Conference on Artificial Intelligence and Law (ICAIL 2001)*, New York, ACM, 2001

[14] BOER, HOEKSTRA and WINKELS, *Metalex: Legislation in XML*, *op. cit.* (as in n. **??**), p. 1.

[15] MOENS, MARIE-FRANCINE, "Innovative techniques for legal text retrieval", *Artificial Intelligence and Law*, 9 2001

[16] TURTLE, H., "Text retrieval in the legal world", *Artificial Intelligence and Law*, 3 1995

[17] BOER, HOEKSTRA and WINKELS, *Metalex: Legislation in XML*, *op. cit.* (as in n. **??**), p. 3.

[18] BOER, HOEKSTRA and WINKELS, *Metalex: Legislation in XML*, *op. cit.* (as in n. **??**), p. 3.

> "The MetaLex standard supports multi-lingual documents in two distinct ways: through localization of XML elements and by providing the means to maintain multiple language versions of the same document in one file."[19]
>
> "To keep track of versions MetaLex provides a number of attributes for every structural XML element in the document that can be identified, selected, and thus changed; the date-publication of an element is the time the element is officially published or announced. The date-enacted, the time the content becomes applicable in decisionmaking, is always later than or the same as date-publication, but before date-repealed, the time the content becomes inapplicable in decisionmaking. Between date-enacted and date-repealed the element and its content is active, and outside this interval it is inactive. Table 1 can be used to deduce active time intervals from the presence or absence of these attributes. The date-version attribute represents the date the correctness of the content and other dates of the XML element was last verified. The XML document looses its value as a normative reference as time progresses and the time-interval between date-version and today increases."[20]

MetaLex is extensible in several ways:[21]

- language extensions
- information about texts can be added in RDF statements

The work on MetaLex is being continued in the Estrella Project `http://www.estrellaproject.org/` The main technical objectives of the Estrella project are to develop a Legal Knowledge Interchange Format (LKIF), building upon emerging XML-based standards of the Semantic Web.[22] LKIF uses MetaLex as a standard for sources of law, and participates in the MetaLex CEN workshop to align MetaLex and LKIF. LKIF is used to represent the meaning of sources of law, for the purpose of building knowledge based systems. In MetaLex terms, it is a schema for MetaLex metadata.

[19] BOER, HOEKSTRA and WINKELS, *Metalex: Legislation in XML*, *op. cit.* (as in n. **??**), p. 6.
[20] BOER, HOEKSTRA and WINKELS, *Metalex: Legislation in XML*, *op. cit.* (as in n. **??**), p. 5.
[21] BOER, HOEKSTRA and WINKELS, *Metalex: Legislation in XML*, *op. cit.* (as in n. **??**), p. 7.
[22] HOEKSTRA, RINKE et al., "The LKIF Core Ontology of Basic Legal Concepts", In CASANOVAS, POMPEU et al. (ed.), *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*, Stanford, CA, USA,, 2007 , `http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-321/`.

The MetaLex schema has been incorporated into a CEN Workshop Agreement[23] A CWA is accepted by the CEN and associated standard organisations as a publicly available specification (PAS or pre-norm) for the period of three years, after which the agreement must be renewed or upgraded to a norm.[24]

> "The MetaLex/CEN schema is based on best practices from amongst others the previous versions of the MetaLex schema, the Akoma Ntoso schema, and the Norme in Rete schema. Other important sources of inspiration are i.a. LexDania, CHLexML, FORMEX, R4eGov, etc. In addition to these government or open standards there are many XML languages for publishing legislation in use by publishers. Standards like PRISM, in which major publishers are involved, are also a source of inspiration."[25]

Further reading

- Metalex `http://www.metalex.eu/` project page
- VAN ENGERS, T. et al., "POWER: Using UML/OCL for Modeling Legislation -an application report", In X. (ed.), *Proceedings of the 8th International Conference on Artificial Intelligence and Law (ICAIL 2001)*, New York, ACM, 2001, *op. cit.* (as in n. **??**)
- BOER, ALEXANDER, HOEKSTRA, RINKE and WINKELS, RADBOUD, "Metalex: Legislation in XML", In BENCH-CAPON, TREVOR, DASKALOPULU, ASPASSIA and WINKELS, RADBOUD (ed.), *Legal Knowledge and Information Systems: JURIX 2002*, IOS Press, 2002 , `http://www.jurix.nl/pdf/j02-01.pdf`, *op. cit.* (as in n. **??**)
- HOEKSTRA, RINKE et al., "The LKIF Core Ontology of Basic Legal Concepts", In CASANOVAS, POMPEU et al. (ed.), *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*, Stanford, CA, USA,, 2007 , `http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-321/`, *op. cit.* (as in n. **??**)

Other

INT

[23] A CEN Workshop Agreement is a consensus-based specifications, drawn up in an open Workshop environment. `https://www.cen.eu/CENORM/sectors/technicalcommitteesworkshops/workshops/cwas.asp`

[24] `http://www.metalex.eu/about/`

[25] `http://www.cen.eu/cenorm/businessdomains/businessdomains/isss/activity/ws_metalex.asp`.

Akoma Ntoso

Akoma Ntoso `http://www.akomantoso.org/` builds on the research of the Norme in Rete project.

See VITALI, FABIO and ZENI, FLAVIO, "Towards a country-independent data format: the Akoma Ntoso experience", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art5.pdf`

TEI

The Text Encoding Initiative (TEI) SGML DTD's has some extensions which identify some basic structure elements in legal texts.

See FINKE, NICHOLAS D., "TEI Extensions for Legal Text", In X. (ed.), *Proceedings of the Text Encoding Initiative Tenth Anniversary User Conference*, 1997 , `http://xml.coverpages.org/finkeTEI10.html`

EU

FORMEX

Formex `http://formex.publications.europa.eu/index.html` describes the format for the exchange of data between the Publication Office and its contractors. In particular, it defines the logical markup for documents which are published in the different series of the Offical Journal of the European Union.

Corpus Legis

The Corpus Legis `http://www.juridicum.su.se/iri/corpus/index.htm` project developed an SGML DTD for legal documents for the purpose of a large database, it identifies some basic structure elements in legal texts.

See MAGNUSSON SJÖBERG, CECILIA, Stockholm, Jure, 1998

AU

Australian Justice Sector Metadata Scheme (JSMS)

> "The Justice Sector Metadata Standard sets out the metadata elements required to allow consistent retrieval of legal information by search facilities. The standard caters for legal specific information such as jurisdiction, act name and the distinction between primary and secondary materials (legislation vs guides to the law).

> The Justice Sector Metadata Standard arose from the work of the Legal Information Standards Council (LISC) and the LawZone trial (LawZone: A new way of searching, improving community access to legal information on the Internet 1999)."[26]

CH

CHLexML

CHLexML `http://www.chlexml.ch/` is an XML schema for legislation issued by the federal, cantonal and municipal levels.

DK

Lex Dania

Lex Dania XML is a joint project between the Danish Parliament (The Folketing) and the Danish Ministry of Justice, Civilstyrelsen, Retsinformation. Its aim is to develop a standardized XML format for the documents contained in Retsinformation (the Danish central register for Laws and rules), conforming to the Danish Governmental XML-standard.

See PETERSEN, KNUD ERIK, "Lex Dania XMl status april 2005", *I Quaderni*, 2005, Nr. 18 , `http://www.cnipa.gov.it/site/_files/Quaderno\%2018.pdf`

2.1.2 Cross-border and multilingual legal information systems

Advanced legal databases with a national scope are useful tools, however in today's political and economical landscape they alone can provide no more than a partial view of all relevant norms. A comprehensive legal information system would ideally allow to move seamlessly from the national to the the transnational and/or international level and vice versa. Such a system requires the development of an appropriate method to deal with the use of multiple languages.

In the EU, the need for such a system is felt by citizens, businesses and government alike. At present, there is no comprehensive database linking EU directives to the implementing measures taken by the Member States. As a consequence, finding the equivalents of norms transposing directives in the various Member States is a manual task which is very time-consuming.

Eulegis

Eulegis `http://canada.esat.kuleuven.be/docarchwebsite/show.jsp?page=projects\&id=EULEGIS` stands for European User Views to Legislative Information in Structured Form.

[26] `http://info.lawaccess.nsw.gov.au/lawaccess/lawaccess.nsf/pages/jsms`

The project aimed to develop single point of access to all data systems providing legislative information in the European Union.[27]

> "Legal information in Europe is scattered in numerous heterogeneous databases. The data in the databases is structured, organized and classified in various ways, the contents are written in different languages, and the retrieval techniques vary. Providing integrated access to the databases would serve both legal experts and laymen. Issues related to the Web access of European legal databases were studied in the EULEGIS project. Requirements for the integrated service were investigated and a prototype system was implemented. The implementation was based on the idea of rich metadata. An XML-based model for the metadata was developed and implemented. The model included data about legal processes, organizational actors in the processes, types of documents created in the processes, and databases providing access to the documents of the types. An important subset of the metadata was visualized in the user interface graphically."[28]

EULEX 3 and Nat-Lex

The research done in Eulegis has been continued in the EULEX 3 and Nat-Lex projects. EULEX 3 aimed to develop an access service for national implementing measures of EU law. There do not appear to be any published results of this project, however the existing database of EU law, Eur-Lex, has been enriched with reference information about national implementing measures. At present, no direct link to the national legislative databases is provided.

The goal of Nat-Lex was to develop a single point of access (with a standard search interface) to legal online information services in EU Member States.[29]

The results of Nat-Lex are a common access portal for national law (N-Lex), which is currently available as an experimental application in 2006. The portal allows users to search national sites using a single uniform search template and provides a multilingual thesaurus. The search possibilities and the results are completely dependent on the national sites, which explains the differences in the availability of search criteria and in the presentation and scope of results.[30]

[27] HIETANEN, AKI, "Networking European Legal Sites : Experiences and Challenges", In X. (ed.), *Proceedings of the Law via the Internet Conference*, Paris, 2004 , `http://www.frlii.org/IMG/pdf/hietanenparis.pdf`.

[28] LYYTIKÄINEN, VIRPI, TIITINEN, PASI T. and SALMINEN, AIRI, "XML Metadata for Accessing Heterogeneous Legal Databases", In X. (ed.), *Proceedings of the XML Europe 2001 Conference*, 2001 , `http://www.gca.org/papers/xmleurope2001/papers/html/s27-4.html`

[29] HIETANEN, *Networking European Legal Sites : Experiences and Challenges*, *op. cit.* (as in n. **??**).

[30] `http://eur-lex.europa.eu/n-lex/pays.html?lang=en`

Globalex

> "GlobaLex http://www.nyulawglobal.org/Globalex/about.htm is an electronic legal publication dedicated to international and foreign law research. Published by the Hauser Global Law School Program at NYU School of Law. GlobaLex is committed to the dissemination of high-level international, foreign, and comparative law research tools in order to accommodate the needs of an increasingly global educational and practicing legal world.
>
> The information and articles published by GlobaLex represent both research and teaching resources used by legal academics, practitioners and other specialists around the world who are active either in foreign, international, and comparative law research or those focusing on their own domestic law. The guides and articles published are written by scholars well known in their respective fields and are recommended as a legal resource by universities, library schools, and legal training courses. The tools available in GlobaLex will continue to expand to cover international law topics, countries and legal systems thus providing a coherent and encompassing research tool for all constituencies."[31]

Further reading

- KUEHL FROSTESTAD, HEIDI, "Globalex: A Unique and Valuable Tool for Foreign, Comparative, and International Law Research", *International Journal of Legal Information*, 34 2006 , http://www.heinonline.org/HOL/Page?handle=hein.journals/ijli34\id=1\size=2\collection=journals\index

2.1.3 Bridging the gap: tying legislation to its application domain

Being able to point to an exact rule or provision in legal source material is one thing. Beyond that, the question remains what the provision really means, which is usually if not always open for interpretation. Realizing this fact, the legislative XML research community has turned to methods for linking legal sources to documents about their interpretation.[32]

The problem of linking legislative knowledge to external (domain-specific) knowledge is also one encountered by law-makers, whose goal is to have an impact on the world around them,

[31] http://www.nyulawglobal.org/Globalex/about.htm

[32] BIAGIOLI, CARLO, "How to link (external) models or interpretations of the meaning of sources of law to the original sources", Leiden, 2007 , http://www.lri.jur.uva.nl/~winkels/PP-Jurix-2007.pdf.

and not to write legislation per se. Therefor, the research in this area[33] can be of interest in the inverse situation where a pre-existing legal text needs to be applied to a specific situation.

Further reading

- BIAGIOLI, CARLO, "How to link (external) models or interpretations of the meaning of sources of law to the original sources", Leiden, 2007 , `http://www.lri.jur.uva.nl/~winkels/PP-Jurix-2007.pdf`, *op. cit.* (as in n. **??**)
- BIAGIOLI, CARLO et al., "Law Making Environment. Perspectives", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art19.pdf`, *op. cit.* (as in n. **??**)

2.2 Legal semantic web

Online databases of regulatory texts are a great tool for human users, however information technology has much more to offer us in terms of support of the practice of law. The next step is to enhance the natural language legal texts with computer readable specifications, allowing for all kinds of automated processing, such as document retrieval, accessing related information, establishing consolidated version and even applying the rules contained in the text. Researchers are turning to semantic web technologies, which process information according to its content or meaning, to realize these advances.[34]

Semantic web is not a single technology, but builds upon a number of different technologies:

- XML is used to separate natural language elements from computer-readable elements.
- RDF is a specification for subject-predicate-object expressions (e.g. the sky has the color blue) which is used to describe properties of an entity or resource.
- Ontologies provide a description of entities that exist in a domain and the relationships between them.

Another – still experimental – component of semantic web technology are rule languages and rule inference mechanisms. Obviously, extending semantic web with legal logic is of great interest for legal informatics researchers.

[33] See BIAGIOLI, CARLO et al., "Law Making Environment. Perspectives", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art19.pdf`.

[34] `http://www.one-lex.eu/The%20Project/Mission/Mission.html`

2.2.1 Legal RDF

Developmant of an RDF 'dictionary' for the legal domain was the focus of research in the LeXML initiative. The aim was to identify and describe similarities and differences between legal concepts in different languages so that legal documents built up with different DTDs could be exchanged automatically.[35]

LEXML Network members

- OAISIS Member Section LegalXML `http://legalxml.org/`
- Germany `http://www.lexml.de/`
- Sweden `http://www.juridicum.su.se/lexml/lexml.htm`
- Italy `http://www.lexml.it/`
- The Netherlands `http://law.leiden.edu/xml/` (no longer available, see Internet Archive `http://web.archive.org/web/20061125172713/http://law.leiden.edu/xml/`)
- Spain `http://www.uv.es/lexml/`
- U.S. `http://www.legalxml.org`

There appears to have been little development of the Legal RDF dictionary in recent years.

2.2.2 Legal ontologies

Much work is being done on the development of legal ontologies. Legal ontologies describe which concepts are used in the legal field in a machine-readable way. Therefor ontologies form a building block for technologies that aim to translate legal texts, policies and legal knowledge in general into machine-readable form.

Further reading

- RYAN, HENRY et al., "Ontology-Based Platform for Trusted Regulatory Compliance Services", In X. (ed.), *On The Move to Meaningful Internet Systems 2003: OTM 2003 Workshops*, Volume 2889, Lecture Notes in Computer Science, Berlin, Springer, 2003 , `http://dx.doi.org/10.1007/b94345`

[35] MULLER, MURK, "Legal RDF Dictionary", In X. (ed.), *Proceedings of XML Europe 2002*, 2002 , `http://www.idealliance.org/papers/xmle02/dx_xmle02/papers/03-04-03/03-04-03.html`

- KABILAN, VANDANA, JOHANNESSON, PAUL and RUGAIMUKAMU, DICKSON M., "Business Contract Obligation Monitoring through Use of Multi Tier Contract Ontology", In X. (ed.), *On The Move to Meaningful Internet Systems 2003: OTM 2003 Workshops*, Volume 2889, Lecture Notes in Computer Science, Berlin, Springer, 2003 , http://dx.doi.org/10.1007/b94345
- LEHMANN, JOS et al. (ed.), *LOAIT - Legal Ontologies and Artificial Intelligence Techniques*, Volume 4, IAAIL Workshop Series, Tilburg, Wolf Legal Publishers, 2005
- CASANOVAS, POMPEU et al. (ed.), Stanford, CA, USA,, Stanford University, 2007, http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-321/
- RUBINO, ROSSELLA, ROTOLO, ANTONINO and SARTOR, GIOVANNI, "An OWL Ontology of Norms and Normative Judgements", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art13.pdf

JurWordnet

Natural languages are difficult to master by machines for an number of reasons. The exact meaning of a term is often hard to define and quite often shifts over time and according to context. Terms may have more than one meaning, and one meaning can be covered by many synonymous terms. To tackle this particular problem WordNet http://wordnet.princeton.edu/ was developed at Princeton University.

> "WordNet® is a large lexical database of English, developed under the direction of George A. Miller. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts can be navigated with the browser. WordNet is also freely and publicly available for download. WordNet's structure makes it a useful tool for computational linguistics and natural language processing."[36]

Legal terminology is based in common language, however in the legal field common words often get a more specific meaning. The JurWordnet http://www.ittig.cnr.it/Ricerca/

[36] http://wordnet.princeton.edu/

UnitaEng.php?Id=11\&T=4 project aims to create a WordNet for the legal field which is connected to the generic WordNet.

WordNets can be used to support information retrieval technology or for automatic production of metadata.[37]

Further reading

- MILLER, GEORGE A., "WordNet: A lexical database for english", *Communications of the ACM*, vol. 38 1995, Nr. 11
- VOSSEN, PIEK, *EuroWordNet A Multilingual Database with Lexical Semantic Networks*, Dordrecht, Kluwer Academic Publishers, 1998
- FELLBAUM, CHRISTIANE, *WordNet: An electronic lexical database*, Cambridge, Mass., MIT Press, 1998

LOIS

An obvious next step, though challenging to realize, is the creation of multilingual WordNets. This was the aim of the LOIS project:

> "The main objective of LOIS http://www.ittig.cnr.it/Ricerca/UnitaEng.php?Id=70\&T=4 is the localization of WordNets describing the legal domain into 6 different European languages, namely Italian, English, German, Czech, Portuguese and Dutch. The synsets (or concepts) of these WordNets will be linked across them, in such a way to guarantee cross lingual access to European legislation and other legal documents (such as court cases). The citizen and/or the professional user will then be enabled to enter queries to a legal documentation base into his/her language and retrieve also documents written in different languages."[38]

Further reading

- PETERS, WIM, SAGRI, MARIA-TERESA and TISCORNIA, DANIELA, "The structuring of legal knowledge in LOIS", *Artificial Intelligence and Law*, Vol. 15 2007, Nr. 2

[37] http://www.ittig.cnr.it/Ricerca/materiali/JurWordNet/UsingJurWordNet.htm
[38] LOIS http://www.ittig.cnr.it/Ricerca/UnitaEng.php?Id=70\&T=4

- TISCORNIA, DANIELA, "The Lois Project: Lexical Ontologies for Legal Information Sharing", In BIAGIOLI, CARLO, FRANCESCONI, ENRICO and SARTOR, GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art14.pdf
- TISCORNIA, DANIELA, "Metadata for Content Description", *I Quaderni*, 2005, Nr. 18 , http://www.cnipa.gov.it/site/_files/Quaderno\%2018.pdf
- AJANI, GIANMARIA et al., "Multilingual Conceptual Dictionaries Based on Ontologies", In BIAGIOLI, CARLO, FRANCESCONI ENRICO SARTOR GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art12.pdf

IPronto: ontology for IPR

The ultimate goal of developing legal ontologies is to allow for automated processing of information with legal significance. A prime candidate for early adoption of such technology is the intellectual property domain. The rules governing intellectual property, copyright in particular, are harmonized worldwide to a high degree. The content industry has a great interest in automatically controlling use of their intellectual property, as can be seen in the large investements already done in copy and use controlling techniques. User resistence to such techniques aside, a major criticism of current control mechanisms is precisely that they do not follow legal boundaries, notably where the application of copyright exemptions is concerned. The use of legal ontologies hold much promise in this respect.

> "Ontologies are so expressive that they will be able to capture a great amount of the underlying legal framework and combine it with the usage models typical in RELs initiatives. Therefore, it will be possible to develop a copyright ontology that takes into account copyright law together with the common usage patterns of copyrighted content.
>
> Ontologies have the additional benefit of facilitating evolvability and interoperability. Therefore, a copyright ontology can be defined with the required level of detail for a given application context and evolve later in order to cope with new situations and requirements. And these new requirements are going to appear for sure due to the dynamism of digital technologies and global networks and markets.
>
> On the other hand, it is quite unlikely that there is going to be just a one-fits-all solution for rights expressions representation. Therefore, interoperability is going

> to be a key issue and ontologies an opportunity. As ontologies do not constrain the way things are written down, i.e. the grammar, but just what are we talking about, i.e. the semantics, it is easier to interoperate. A copyright ontology will thus also facilitate interoperability among different RELs. Moreover, it will be easily enriched with the semantics that will be reused from existing initiatives, which will facilitate the development of the copyright ontology, its validation and enable it as a key tool for DRM interoperability and integration. "[39]

The IPronto `http://dmag.upf.edu/ontologies/ipronto/` project builds upon web ontologies "to facilitate the automation and interoperability of IPR frameworks integrating both parts, called Rights Expression Language and Rights Data Dictionary. These objectives can be accomplished using ontologies, that can provide the required definitions of the rights expression language terms in a machine-readable form. Thus, from the automatic processing point of view, a more complete vision of the application domain is available and more sophisticated processes can be carried out. Moreover, the modularity of web ontologies, constituted by concept and relation definitions openly referenceable, allows their free extension and adaptation."[40]

IPronto is not being developed in isolation of other initiatives, but builds on previous work.[41] The work started in IPronto is being continued in the NewMars `http://dmag.upf.edu/newmars/` project, which aims to "develop a multimedia contents e-Commerce platform that manages the Intellectual Property Rights associated to them."[42]

Further reading

- DELGADO, JAIME et al., "IPROnto - Intellectual Property Rights Ontology", *ISWC*, 2002 , `http://dmag.upf.edu/ontologies/ipronto/ISWCPoster.pdf`
- DELGADO, JAIME et al., "IPROnto: An Ontology for Digital Rights Management", In BOUNCIER, D. (ed.), *Legal Knowledge and Information Systems*, Jurix 2003, Amsterdam, IOS Press, 2003 , `http://www.jurix.nl/`, *op. cit.* (as in n. **??**)

[39] GARCÍA, ROBERTO, *A Semantic Web Approach to Digital Rights Management*, Barcelona, Spain, Department of Technologies, Universitat Pompeu Fabra, 2005 , `http://rhizomik.net/~roberto/thesis/`, p. 106. In his thesis paper, the author models a single legal system as it is at a specific point in time. How to deal with changes of the rules or addresses differences amongst legal systems is outside the work's scope.

[40] `http://dmag.upf.edu/ontologies/ipronto/`

[41] DELGADO, JAIME et al., "IPROnto: An Ontology for Digital Rights Management", In BOUNCIER, D. (ed.), *Legal Knowledge and Information Systems*, Jurix 2003, Amsterdam, IOS Press, 2003 , `http://www.jurix.nl/`.

[42] `http://dmag.upf.edu/newmars/`

- GARCÍA, ROBERTO, GIL, ROSA and DELGADO, JAIME, "Intellectual Property Rights Management using a Semantic Web Information System", In X. (ed.), *OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2004*, Lecture Notes in Computer Science, Berlin, Springer, 2004, 3291
- GARCÍA, ROBERTO, *A Semantic Web Approach to Digital Rights Management*, Barcelona, Spain, Department of Technologies, Universitat Pompeu Fabra, 2005 , `http://rhizomik.net/~roberto/thesis/`, *op. cit.* (as in n. **??**)
- GARCÍA, ROBERTO, GIL, ROSA and DELGADO, JAIME, "A web ontologies framework for digital rights management", *Artificial Intelligence and Law*, Vol. 15 2007, Nr. 2 , `http://dx.doi.org/10.1007/s10506-007-9032-6`

2.2.3 Legal Logic

Descriptions of concepts and the links between them alone are not enough to allow for automated processing of legal information. A method to express legal rules or legal logic in machine-readable form is required. This is being researched in a number of projects.

Further reading

- VAN ENGERS, T. et al., "POWER: Using UML/OCL for Modeling Legislation -an application report", In X. (ed.), *Proceedings of the 8th International Conference on Artificial Intelligence and Law (ICAIL 2001)*, New York, ACM, 2001
- KERRIGAN, SHAWN and LAW, KINCHO H., "Logic-based regulation compliance-assistance", New York, ACM Press, 2003, 126–135

RuleML

RuleML `http://ruleml.org/` is an XML based language for the representation of rules. It offers facilities to specify different types of rules from derivation rules to transformation rules to reaction rules. RuleML is developed by the Rule Markup Initiative, which aims to develop RuleML as the canonical Web language for rules using XML markup, formal semantics, and efficient implementations.[43]

In their paper 'Modelling Contracts Using RuleML', Governatori and Rotolo use RuleML to develop a conceptual representation of contracts. From a contract in natural language they distill

[43] See RuleML Mission Statement, `http://ruleml.org`

the obligations and rules in a computer-readable form, thus allowing for automated monitoring.[44]

> "Business contracts are mutual agreements between two or more parties engaging in various types of economic exchanges and transactions. They are used to specify the obligations, permissions and prohibitions that the signatories should hold responsible for and to state the actions or penalties that may be taken when any of the stated agreements is not being met. Given the increasing efforts by organisations to carry out their business via the Internet, it is crucial to model contracts in terms of workflows, so that all relevant tasks of contracts can be described as business processes, where business processes are defined by business rules, statements or policies listed in business contracts or other legal documents that are used by organisations to run the activities, to provide an understanding of how a business operates, and to direct the behaviour of the organisation."[45]

From monitoring contract execution to compliance with internal policies is just a small step. Exactly how RuleML could be used to support compliance monitoring is beyond the scope of this report.

REALM: Regulations Expressed As Logical Models

REALM `http://www.zurich.ibm.com/csc/security/compliance.html` is a project that has automated support of compliances as its main goal:

> "Recent years have seen a number of high-profile incidents ofcorporate accounting fraud, security violations, terrorist acts, and disruptions of major financial markets. This has lead to a proliferation ofnew regulations that directly impact businesses. As a result, businesses,in particular publicly traded companies, face the daunting task of complying with an increasing number of intricate and constantly evolvingregulations. Together with the growing complexity of today's enterprisesthis requires a holistic compliance management approach with the goalof continually increasing automation.We introduce REALM (Regulations Expressed as Logical Models), a metamodel and method for modeling regulations and managing them ina systematic

[44] GOVERNATORI, GUIDO and ROTOLO, ANTONINO, "Modelling Contracts Using RuleML", In GORDON, T. (ed.), *Legal Knowledge and Information Systems*, Jurix 2004, Amsterdam, IOS Press, 2004 , `http://www.jurix.nl/pdf/j04-16.pdf`.

[45] GOVERNATORI and ROTOLO, *Modelling Contracts Using RuleML*, *op. cit.* (as in n. **??**), p. 141.

> lifecycle in an enterprise. We formalize regulatory requirements as sets of compliance rules in a novel real-time temporal objectlogic over concept models in UML, together with metadata for traceability. REALM provides the basis for subsequent model transformations,deployment, and continuous monitoring and enforcement of compliancein real business processes and IT systems."[46]

Further reading

- GIBLIN, C. et al., "Regulations Expressed as Logical Models (REALM)", In MOENS, MARIE-FRANCINE and SPYNS, PETER (ed.), *Proceedings of the 18th Annual Conference on Legal Knowledge and Information Systems*, Brussels, Jurix, 2005, *op. cit.* (as in n. **??**)
- PFITZMANN, BIRGIT, POWERS, CALVIN and WAIDNER, MICHAEL, *IBM's Unified Governance Framework (UGF) Initiative*, Zurich, IBM Research Division, 2007, IBM Research Report RZ 3699 , `http://www.zurich.ibm.com/pdf/csc/rz3699_UGF-whitepaper.pdf`
- IBM Governance & Compliance `http://www.zurich.ibm.com/csc/security/compliance.html`.

ALIS project

One of the ALIS project `http://www.alisproject.eu` objectives is to support regulatory compliance in the field of intellectual property rights.

> "ALIS will check whether actions and decisions of Government Departments are compliant with the applicable legal and regulatory framework, by resorting to information technologies, knowledge representation, computational logics and formal reasoning. The same technologies may also serve individuals and private companies.
>
> Furthermore, ALIS will analyze compliance with respect to different systems of laws and regulations, thus pointing out the potential contradictions that may occur when several systems of laws and regulations need to be taken into account simultaneously.
>
> The benefits of regulatory compliance for the actors under consideration are :

[46] GIBLIN, C. et al., "Regulations Expressed as Logical Models (REALM)", In MOENS, MARIE-FRANCINE and SPYNS, PETER (ed.), *Proceedings of the 18th Annual Conference on Legal Knowledge and Information Systems*, Brussels, Jurix, 2005, p. 1.

1. Minimization of legal risk, and hence of possibly high expenses to cover these risks
2. Cost reduction at the level of the judicial system
3. Better governance at the governmental level
4. Harmonisation of best practices at the European level."[47]

Further reading

- CEVENINI, CLAUDIA et al., "Development of the ALIS IP Ontology: Merging Legal and Technical Perspectives", In X. (ed.), *Computer-Aided Innovation (CAI), IFIP International Federation for Information Processing*, Boston, Springer, 2008 , http://dx.doi.org/10.1007/978-0-387-09697-1_14

Policy aware web

"Policy awareness is a property of the Web that will provide users with accessible and understandable views of the policies associated with resources, enable agents to act in response to rules on a user's behalf, thereby making compliance with stated rules easier, and afford a greater opportunity for accountability when rules are intentionally or accidentally broken."[48]

There are quite a few initiatives developing new technologies for expressing information policies on the web. Characteristic for these initiatives is that they deal with a specific domain (privacy, copyright, identity or access control), depending on the needs of the community in question. At their core these vertical new technologies deal with the same three entities: people, content and permissions. Iannella e.a. propose that semantic web technology should allow for the development of a policy framework that captures the commonalities between these policy languages whilst supporting the specific needs of each.[49] The main challenges for such a policy framework are how to arrive at an abstract model for the policy framework, how to create representations of the model and further how to then implement this model in concrete applications.[50]

[47] Alis Project Website, Regulatory Compliance, http://www.alisproject.eu/index.php?option=com_content&task=view&id=41&Itemid=68

[48] WEITZNER, DANIEL J. et al., "Transparency and End-to-End Accountability: Requirements for Web Privacy Policy Languages", In X. (ed.), *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, W3C, 2006 , http://www.w3.org/2006/07/privacy-ws/papers/, p. 2.

[49] IANNELLA, RENATO, HENRICKSEN, KAREN and ROBINSON, RICKY, "A Policy Oriented Architecture for the Web: New Infrastructure and New Opportunities", Ispra, Italy, W3C, 2006 , http://www.w3.org/2006/07/privacy-ws/papers/05-ianella-policy-oriented-architecture, p. 2.

[50] IANNELLA, HENRICKSEN and ROBINSON, *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006 2006*, *op. cit.* (as in n. **??**), p. 3.

Further reading

- WEITZNER, DANIEL J. et al., "Transparency and End-to-End Accountability: Requirements for Web Privacy Policy Languages", In X. (ed.), *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, W3C, 2006 , `http://www.w3.org/2006/07/privacy-ws/papers/`, *op. cit.* (as in n. **??**)
- IANNELLA, RENATO, HENRICKSEN, KAREN and ROBINSON, RICKY, "A Policy Oriented Architecture for the Web: New Infrastructure and New Opportunities", Ispra, Italy, W3C, 2006 , `http://www.w3.org/2006/07/privacy-ws/papers/05-ianella-policy-oriented-` *op. cit.* (as in n. **??**)

Chapter 3

Evidence

Simply put, evidence must answer questions that go along the lines of who, what, why, when, where and how. Electronic evidence has challenged existing legal practices with respect to collection, production and evaluation of proof. In the discussion of electronic evidence, two distinct legal domains can be discerned, one is cyber-crime and the other is electronic transactions in a broad sense. Though electronic records management, as a business process, falls within the scope of the latter, insights gained in the former can be instructive.

There appears to be no general consensus on a defintion of electronic evidence. Evidence is considered to be all information by which facts tend to be proved. Evidence is thus the means by which the facts constituating an offence or surrounding a transaction can be proved.[1] Electronic evidence then can be defined as any information obtained from an electronic device or digital medium which serves to convince the truth of a fact or action.[2]

Electronic evidence is not intrinsically different from other types of evidence, however the fragility and the transience of many forms of computer evidence raise additional concerns.[3] In order to guarantee the authenticity of the records it contains, an electronic records management system must implement appropriate information security techniques. Regulatory compliance monitoring and enforcement too is likely to use similar techniques. In this sense, taking a look

[1] See LEROUX, OLIVIER, "Legal admissibility of electronic evidence", *International Review of Law, Computers Technology*, Vol. 18 2004, Nr. 2, p. 196 ff.

[2] Compare with X. (ed.), *The admissibility of electronic evidence in court: fighting against high-tech crime*, Barcelona, Cybex, 2006 , `http://www.cybex.es/agis2005/docs/libro_aeec_en.pdf`

[3] KARYDA, MARIA and MITROU, LILIAN, "Internet forensics: legal and technical issues", *Proceedings of the second International Workshop on Digital Forensics and Incident Analysis* 2007, LEROUX, *International Review of Law, Computers Technology Vol. 18 [2004]*, *op. cit.* (as in n. **??**), EUROPE, RAND (ed.), *Handbook of Legislative Procedures of Computer and Network Misuse in EU Countries – Study for the European Commission*, Directorate-General Information Society, 2002.

at digital forensics seems appropriate.

3.1 Digital Forensics

Traditionally, forensic science or forensics in the broad sense is the use of science to answer legal questions, though the term is most commonly used in the narrower sense of scientific analysis of evidence gathered in criminal investigations.[4] Digital forensics is the analysis of evidence in digital form, though the term is used both for official criminal investigations as for private investigations of wrong-doings.

> "Digital forensic investigations (DFIs) are commonly employed as a post-event response to a serious information security or criminal incident. They typically consider the case when the PC of a suspect has been seized. The hard-drive is imaged and an investigation proceeds to search for traces of evidence. The examination is conducted in a systematic, formalised and legal manner to ensure the admissibility of the evidence. The process of a digital forensic investigation is subject to considerable scrutiny of both the integrity of the evidence [Sommer 1998][5], and the integrity of the investigation process [Stephenson 2002, 2003b][6]."[7]

> "Internet forensics involve the recognition, recovery and reconstruction of digital evidence and its management in a way that renders it admissible in prosecution and – more generally – in legal proceedings."[8]

> "Procedural problems arise from the lack of standardization, as well as the lack of theoretical framework for the field of digital forensics. Using ad-hoc methods and

[4] See Merriam-Webster Dictionary, `http://www.m-w.com`.

[5] SOMMER, P., "Intrusion Detection Systems as Evidence", In X. (ed.), *Proceedings of Recent Advances in Intrusion Detection 1998*, 1998 , `http://www.raid-symposium.org/raid98/Prog_RAID98/Full_Papers/Sommer_text.pdf`

[6] STEPHENSON, P., "A Comprehensive Approach to Digital Incident Investigation", *Information Security Technical Report*, Vol. 8 2003, Nr. 2 , `http://dx.doi.org/10.1016/S1363-4127(03)00206-1`; STEPHENSON, P., "Using Evidence Effectively", *Computer Fraud and Security*, 2003, Nr. 3 , `http://dx.doi.org/10.1016/S1361-3723(03)03012-4`

[7] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]* (as in n. **??**), p. 1.

[8] KARYDA and MITROU, *Proceedings of the second International Workshop on Digital Forensics and Incident Analysis 2007* (as in n. **??**), referring to MITRAKAS, ANREAS, ZAITCH DAMIEN, "Law, Cybercrime and digital forensics: Trailing Digital Suspects", In KANELIS, PANAGIOTIS, KIOUNTOUZIS EVANGELOS KOLOKOTRONIS NICHOLAS DRAKOULIS MARTAKOS (ed.), *Digital Crime and Forensic Science in Cyberspace*, London, Idea Group, 2006.

> tools for the elicitation of digital evidence can limit the reliability and credibility of the evidence, especially in a crime prosecution process where both the evidence and the processes used for collecting it can be disputed."[9]

Once an incident has occurred and an investigation has to be undertaken, it is too late to ensure that systems are able to provide sufficient and relieble evidence of the facts. Organisations must prepare for such events beforehand to keep the cost of investigations within reasonable limits, yet still be in a position to procure effective evidence when needed. Rowlingson calls this 'digital forensics readiness' and proposes a ten step process to achieve this.[10]

Further Reading

- AHMAD, A., "The Forensic Chain of Evidence Model: Improving the Process of Evidence Collection in Incident Handling Procedures", In X. (ed.), *Proceedings of the 6th Pacific Asia Conference on Information Systems*, 2002 , http://www.dis.unimelb.edu.au/staff/atif/AhmadPACIS.pdf
- STEPHENSON, P., "End-to-End Digital Forensics", *Computer Fraud and Security*, 2002, Nr. 9 , http://dx.doi.org/10.1016/S1361-3723(02)00914-4
- STEPHENSON, P., "Using Evidence Effectively", *Computer Fraud and Security*, 2003, Nr. 3 , http://dx.doi.org/10.1016/S1361-3723(03)03012-4, *op. cit.* (as in n. **??**)
- WOLFE, HENRY B., "Evidence Analysis", *Computers Security*, Vol. 22 2003, Nr. 4 , http://dx.doi.org/10.1016/S0167-4048(03)00404-8
- ROWLINGSON, ROBERT, "A Ten Step Process for Forensic Readiness", *International Journal of Digital Evidence*, vol. 2 2004, Nr. 3 , http://www.utica.edu/academic/institutes/ecii/publications/articles/A0B13342-B4E0-1F6A-156F501C49CF5, *op. cit.* (as in n. **??**)
- ENFSI (ed.), *Guidelines for Best Practice in the Forensic Examination of Digital Technology*, ENFSI, 2006 , http://www.enfsi.eu/uploads/files/ENFSI_Forensic_IT_Best_Practice_GUIDE_5\%5B1\%5D.0.pdf
- MITRAKAS, ANREAS, ZAITCH DAMIEN, "Law, Cybercrime and digital forensics: Trailing Digital Suspects", In KANELIS, PANAGIOTIS, KIOUNTOUZIS EVANGELOS KOLOKOTRONIS NICHOLAS DRAKOULIS MARTAKOS (ed.), *Digital Crime and Forensic Science in Cyberspace*, London, Idea Group, 2006, *op. cit.* (as in n. **??**)

[9] KARYDA and MITROU, *Proceedings of the second International Workshop on Digital Forensics and Incident Analysis 2007*, *op. cit.* (as in n. **??**).

[10] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**).

- KARYDA, MARIA and MITROU, LILIAN, "Internet forensics: legal and technical issues", *Proceedings of the second International Workshop on Digital Forensics and Incident Analysis* 2007, *op. cit.* (as in n. **??**)

3.1.1 Digital archiving readiness

Thus, where forensic readiness – In the context of enterprise security – can be defined as "the ability of an organisation to maximise its potential to use digital evidence whilst minimising the costs of an investigation"[11], similarly, digital archiving readiness could be defined as "the ability of an organisation to maximise its potential to use digital records whilst minimising the costs of proving their authenticity".

Aims[12]:

- To gather admissible evidence legally and in the normal course of business processes;
- To gather evidence targeting the potential disputes that may adversely impact an organisation;
- To ensure that evidence makes a positive impact on the outcome of any legal action;

Steps[13]:

- Define the business scenarios that require digital evidence:
- Identify available sources and different types of potential evidence.
- Determine the evidence collection requirement.
- Establish a capability for securely gathering legally admissible evidence to meet the requirement.
- Establish a policy for secure storage and handling of potential evidence.
- Train staff in digital archiving practices, so that all those involved understand their role in the digital evidence process and the legal sensitivities of evidence.

Define the business scenarios that require digital evidence

Predicting what kind of evidence one may need in the future can be very difficult to do. However, haphazardly recording information about some parts of business processes and not others is a very risky strategy. It makes more sense to analyse the organisations business processes to find

[11] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 5.

[12] Analogues to some of the aims of digital forensic readiness, as viewed by Rowlingson, ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 9.

[13] Analogues to some of the steps towards digital forensic readiness, as listed by Rowlingson, ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 9.

out what (digital) evidence they require, either because it is a legal requirement, a contractual one or an internal one.[14]

If afterwards the analysis proves to have been insufficient, for instance through a failure to procure evidence in court, the plan can be adjusted for the future.

Ideally, for all the records being preserved by an organisation, it should be known why they are created and preserved, e.g. demonstrating compliance with regulations, evidence of contractual agreements, ...

At the same time, the organisation should be confident that records are being created and preserved to fulfill all evidence needs.

As far as documentation goes, there is interest in linking records management policies to the evidence requirements it fulfills. Whenever the requirements change, e.g. laws are modified or contracts amended, the relevant policies can be changed as well.

Identify available sources and different types of potential evidence

Some basic questions need to be asked about possible evidence sources, including[15]:

- Where is data generated?
- What format is it in?
- For how long is it stored?
- How is it currently controlled, secured and managed?
- Who has access to the data?
- How much is produced?
- Is it archived? If so where and for how long?
- How much is reviewed?
- What additional evidence sources could be enabled?
- Who is responsible for this data?
- Who is the formal owner of the data?
- How could it be made available to an investigation?
- To what business processes does it relate?
- Does it contain personal information?

> "Email is an obvious example of a potential rich source of evidence that needs careful consideration in terms of storage, archiving, auditing, and retrieval. But this

[14] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 10.
[15] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 11.

> is not the only means of communication used over the Internet. There is also instant messaging, web-based email that bypasses corporate email servers, chat rooms and newsgroups, and even voice over the Internet. Each of these may need preserving and archiving. A worst case scenario has some of this traffic encrypted."[16]

Records, as a byproduct of business processses, are one category of evidence. The strength of such evidence often rests on the fact that it has been collected according to standard documented business procedures.[17]

A different source of evidence stems from monitoring or surveillance of user activities. Monitoring is usually cited as a means to deter and/or detect crime[18], however it may also play an important role in enforcement of business policies, for instance with respect to archiving of records and compliance. Monitoring touches upon the right to privacy and other human rights of those subjected to it, thus care must be taken to ensure it is done legally.[19]

Determine the evidence collection requirement

A gap analysis between evidence needs and available sources, as well as a cost benefit analysis comparing the costs of collecting the evidence required and the projected benefits should be conducted.[20]

On the benefit side:

- Can the evidence make an impact on the likely success of any formal action?[21]
- Can the evidence be gathered legally without infringing employee rights[22]

On the cost side[23]:

- Cost of monitoring (including tools and staff-time)
- Cost of secure storage
- Cost of organising potential evidence – by classifying, indexing and preparation
- Cost and implications of retrieval if evidence is demanded by a court

[16] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 11.
[17] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 12.
[18] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 12.
[19] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 12.
[20] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 13.
[21] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 13.
[22] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 13.
[23] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 13.

Establish a capability for securely gathering legally admissible evidence to meet the requirement.

This step makes a start with the implementation of measures for digital archiving readiness.

Firstly, measures must ensure that available evidence is collected from the relevant sources and that it is preserved as an authentic record. A preliminary question is whether the evidence can be collected in a legal manner.[24] Once collected, appropriate security measures are required to ensure integrity of evidence.[25]

Establish a policy for secure storage and handling of potential evidence.

Evidence is generally called upon only after the lapse of a certain amount of time, months or even years after it was collected.

> "The objective of this step is to secure the evidence for the longer term once it has been collected and to facilitate its retrieval if required. It concerns the long-term or off-line storage of information that might be required for evidence at a later date."[26]

> "A policy for secure storage and handling of potential evidence comprises security measures to ensure the authenticity of the data and also procedures to demonstrate that the evidence integrity is preserved whenever it is used, moved, or combined with new evidence. At all times it must be in a tamper-proof (or tamper-evident state). This corresponds to the use of evidence bags in the physical world. Access to the evidence is controlled and anyone requiring an evidence bag must sign it in and sign it back with the contents unchanged. In the parlance of investigators this is known as continuity of evidence (in the UK) and chain of custody (in the US). The chain of custody also includes records of who held, and who had access to, the evidence (for example from swipe control door logs)."[27]

Train staff in digital archiving practices

Records management technologies support business activities, which are still to a large degree human activities. If the people in an organisation do not follow archiving policies, no amount of

[24] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 15.
[25] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 15.
[26] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 16.
[27] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004]*, *op. cit.* (as in n. **??**), p. 17.

sophisticated information technology can rectify this. Training staff is essential so that all those involved understand their role in the digital archiving process, including some of the legal issues concerning evidence.[28]

3.2 Burden of proof

The notion 'burden of proof' signifies two things in a legal dispute:

1. Who must deliver proof to support a claim?
2. Who loses his claim when failing to provide sufficient proof? (= risk of proof)

When the proof consists of records in a wide sense, a major issue is who posesses the relevant records: the claimant, the defendant or a third party. Each of these persons must consider whether they may, may not or must hand over evidence to the court and/or other parties involved, taking into consideration applicable regulatory and contractual obligations.

In Anglo-Saxon legal systems, making available of evidence is regulated under the rules of discovery. Though continental European legal systems generally do not have a comparable set of rules, some form of duty of cooperation in delivery of evidence often exists.

Further reading

- BERGER, CHRISTIAN, "Beweisführung mit elektronischen Dokumenten", *NJW*, 2005, Nr. 15
- MASON, STEPHEN, "Archiving and storing e-mails - The legal and practical issues", *Computer Law Security Report*, vol. 24 2008, Nr. 2 , `http://dx.doi.org/10.1016/j.clsr.2007.09.004`
- The Sedona Conference `http://www.thesedonaconference.org/`

3.3 Admissibility of evidence

Each country has its own rules regarding the admissibility of evidence in court. Common to most countries is that matters of fact may be proven with any available means of evidence. Where there are restrictions as to the form of admissible evidence, these usually only apply to purposeful legal acts (e.g. contracts, declarations, ...). The reason is quite straightforward, facts

[28] ROWLINGSON, *International Journal of Digital Evidence vol. 2 [2004], op. cit.* (as in n. **??**), p. 21.

such as accidental occurances, scientific observations and the acts of others or not subject to the control of the person who must deliver the proof. This person does not have the opportunity to create reliable evidence as the events unfold, unlike in case of intentional acts.

In evidence laws two major systems can be identified, the regulated and the free evidence regime. The regulated evidence regime entails that the law lists which kinds of evidence are permissible as evidence. For instance, the French Civil Code stipulates that only a signed act is sufficient evidence of contracts (except for those representing a low monetary value). Digital files that bear no kind of legal signature, transcripts of chat sessions, etc. may in principle not be presented before the courts. Note that even in more or less strictly regulated evidence regimes, the proof of merely factual matters is open to all kinds of evidence.

In other legal systems, notably the Anglo-Saxon inspired systems, there is no predetermined list of admissible evidence. However, evidence must conform to a number of principles, for instance relevancy or the 'best evidence rule', otherwise it is discarded.[29]

Multinationals face the difficult task of designing their work processes with admissibility requirements of different jurisdictions in mind. If complying with all of them completely proves impossible, choices must be made based on which requirements cost too much to implement in comparison to the projected benefits. It may be cheaper to lose a law suit for lack of sufficient evidence now and again, than to ensure that all evidence is collected and stored for a significant period of time.

Here too, linking the evidence collection and handling policy of an organisation to the legal provisions and principles on which it is based would be a great step forward. For one, if the legal foundations are modified later, the organisation can easily identify which policies may need adaptation. Providing such links may be just as valuable or even more so when policies do not comply with the legal provisions of a country. In such a case the possible consequences of non-compliance can be noted, as well as the reasons for maintaining a non-compliant policy.

3.4 Probative value of digital evidence

Once evidence has been accepted by the court as admissible, the question remains how convincing the judge will find it. Authenticity of evidence is key. Being able to give sufficient assurances about the integrity is a great factor in determining the probative value of records.

[29] About UK evidence rules, see KEARSLEY, AMANDA J., "Electronic Document Management, Legal admissibility of evidence held in electronic form", *Computer Law Security Report*, Vol. 15 1999, Nr. 3. About U.S. rules see KENEALLY, ERIN E., "Digital logs - proof matters", *Digital Investigation*, 2004, Nr. 1 , `http://www.elsevier.com/locate/diin`.

Also the identity of the records – what are they and in which context were they created – is of great importance.

In some instances, the law itself may determine the weight to be attached to evidence through legal presumptions. For instance, the observations made by officers of the police or other official agents may be legally presumed to be true, except in very special circumstances. This is the case in Belgium for instance for the observations a notary makes and records in notarial deeds, e.g. time, date and location of the transaction, identity of the parties, ... As Trusted Third Parties become more common in all kinds of business transactions, their assertions as to the date of transactions, the identity of parties or the integrity of records may one day also benefit from such legal presumptions of truthfulness. This will depend of course on who the TTP is and which guarantees for trustworthiness they present.

Legal presumptions determining probative value are exceptional and only hold in the jurisdiction of their origin. Records and/or assertions contained in them which are presumed truthful in one country, in all likelihood do not have such status in other countries.

> "Trustworthy records are vital to an organization. These records help to improve an organization's operations and aid in reducing its liability and costs. The fundamental purpose of record keeping is to establish solid proof and details of events that have occurred. A trustworthy record management system is, therefore, one that can be relied upon to provide irrefutable evidence of all of the events that have been logged. In other words, trustworthiness has to be established on an end-to-end perspective, from the proper preservation of all of the records to the subsequent delivery of the relevant records to an agent seeking the proof. In this white paper, we show that the current limited focus on storing electronic records in Write-Once-Read-Many (WORM) storage is not adequate to ensure that such records are trustworthy. What is really needed is a process we call fossilization – a holistic approach to storing and managing records that ensures that they are trustworthy. Fossilization is composed of three parts. The first, fossilization of storage, guarantees that all records and their associated metadata are reliably stored and securely protected from any modification. The second, fossilization of discovery, ensures that all preserved records pertinent to an enquiry can be quickly discovered and retrieved. The third, fossilization of delivery, warrants that the exact pertinent records are delivered to the agent and that the records are delivered in an intact form. Because of the extremely high stakes involved in tampering with the records, fossilization must be realized very securely. The essential principles for securely implementing fossilization include 1) raising the barrier to any attack; 2) focusing on end-to-end trust; 3)

limiting what has to be trusted; 4) using a simple, well-defined interface between trusted and untrusted components; and 5) verifying all operations."[30]

- ROSSNAGEL, ALEXANDER, PFITZMANN ANDREAS, "Der Beweiswert von E-Mail", *NJW*, 2003, Nr. 17
- BECKER, ARND, *Elektronische Dokumente als Beweismittel im Zivilprozess*, Frankfurt, Peter Lang, 2004
- HSU, WINDSOR W. and ONG, SHAUCHI, *Fossilization: A Process for Establishing Truly Trustworthy Records*, IBM Almaden Research Center, 2004, IBM Research , `http://domino.research.ibm.com/library/cyberdig.nsf/1e4115aea78b6e7c85256b360066f0d02da1cea05c6c61`, *op. cit.* (as in n. **??**)
- ZHU, QINGBO, HSU WINDSOR W., "Fossilized Index: The Linchpin of Trustworthy Non-Alterable Electronic Records", In X. (ed.), *International Conference on Management of Data archive. Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, New York, ACM Press, 2005 , `http://portal.acm.org/citation.cfm?id=1066157.1066203`
- HOFFMANN, MATHIS, "Der Beweiswert elektronischer Dokumente", *DSWR*, 2006, Nr. 3
- KLEIN, SUSANNE, "Die Beweiskraft elektronischer Verträge. Zur Entwicklung der zivilprozessrechtlichen Vorschriften über die Beweiskraft elektronischer Dokumente", *JurPC Web-Dok.* 2007, Nr. 198

3.4.1 Probative value of (digital) copies

Digital technology has provoked an renewed interest in the probative value of copies of original documents. In first instance, many organizations have sought to replace extensive collections of paper documents with digital copies. More and more, the necessity of converting original digital files into preservation formats and the impact of this conversion on legal and probative value is coming to the forefront.

In many countries, the probative value of copies is dealt with in a piecemeal fashion. In Belgium, for instance, various institutions – both public and private – benefit from a 'copy privilege': the copies of document they make or have made under their authority have the same probative value as the originals.

[30] HSU, WINDSOR W. and ONG, SHAUCHI, *Fossilization: A Process for Establishing Truly Trustworthy Records*, IBM Almaden Research Center, 2004, IBM Research , `http://domino.research.ibm.com/library/cyberdig.nsf/1e4115aea78b6e7c85256b360066f0d4/02da1cea05c6c61`.

Very few countries have a comprehensive regulatory approach in place. One such rare example is the Republic of Slovenia. The Slovenian Protection of Documents and Archives and Archival Institutions Act (PDAAIA)[31] contains provisions regarding the procedure for the conversion of documents (see section III Documents PDAAIA) and the probative value of the resulting copies (see Art. 31 ff. PDAAIA in particular).[32]

3.5 Standards

Though there are a great number of standards that relate to records management, only relatively few address legal admissibility and probative value specifically. The British Standards Institute has issued a number of standards on information and records management, and notably a code of practice addressing admissibility and evidential weight of electronic records was published. The first version of the code of practice dates back to 1996 and was revised in 1999. The current version dates from 2004.

- BSI DISC PD0008:1996 `http://www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000000734901` Code of Practice for Legal Admissibility of Information Stored on Electronic Document Management Systems

- BSI DISC PD0008:1999 `http://www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030001674` Code of Practice for The Legal Admissibility and Evidential Weight of Information Stored Electronically

- BIP 0008-1:2004 `http://www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030104568` Code of Practice For The Legal Admissibility Of Information Stored Electronically and BIP 0009-1 `http://www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030107409` Compliance workbook

The code of practice was extended in 2005 with a section on electronic communication of information and one on the linking of electronic identity to documents.

- BIP 0008-2:2005 `http://www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030132417` Code of practice for legal admissibility and evidential weight

[31] See `http://www.arhiv.gov.si/en/archival_regulations_and_standards/` for an English translation of the act.

[32] See HAJTNIK, TATJANA, "Maintaining legal value of a record throughout their lifecycle", *DLM Forum Meeting, Ljubljana, 8-9 april 2008*, 2008 , `http://dlmforum.typepad.com/Slovenia_Hajtnik.pdf`.

of information communicated electronically and BIP 0009-2:2006 http://www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030146667 Compliance Workbook.

- BIP 0008-3:2005 http://www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030132418 Code of practice for legal admissibility and evidential weight of linking electronic identity to documents and BIP 0009-3:2006 http://www.bsi-global.com/en/Shop/Publication-Detail/?pid=000000000030146651 Compliance Workbook.

The BSI has announced it will issue a revised version in September 2008 after a period of public review. The new standard will address issues relating to the authenticity and integrity of electronic information which could potentially be used as evidence.[33]

3.6 Technological tools for capturing and handling evidence

An ambitious research project to develop technological tools to support the captura and management of digital evidence was the CTOSE project. In the wake of the project a non-profit organisation – the CTOSE foundation – was founded, unfortunately it no longer exists. The project website is partially avaible at the Internet Archive.[34] Possibly, one of the industry partners – Qinetiq [35] – may have incorporated the research results into its products and services.

- The CTOSE process model and guidelines focus on the approach to handling potential electronic evidence, which includes the sub-processes: identification, collection, tamper-free storage, restricted and controlled access, analysis, judicial presentation and documentation of electronic evidence, while taking into consideration the requirements of security, privacy, and due legal process. One of the project's major goals was to identify and extend best practices, bringing together law enforcement agencies and civilian investigators, and carrying this into industry to create a network of experts from all disciplines and sectors.[36]

The CTOSE project deliverables were

[33] BSI Media Release, 2 May 2008, http://www.bsi-global.com/en/About-BSI/News-Room/BSI-News-Content/Disciplines/Information-Management/Legal-admissibility-DPC/

[34] http://www.archive.org.

[35] http://www.qinetiq.com/home/security.html

[36] CTOSE CONSORTIUM, *CTOSE Project Results*, 2003 , http://web.archive.org/web/*hh_/www.ctose.org/ResultsPaperv6.pdf, p. 1.

- CTOSE Methodology Model
- Legal Advisor
- Process Model
- C*CAT – Cyber Crime Advisory Tool
- Forensic Readiness Guidelines
- Forensic Autopsy Tool with XML bindings
- CTOSE Demonstrator
- Project Story Board

CTOSE Methodology Model

The Methodology Architecture Model was used to produce the very detailed investigative Process Model and Forensic Readiness Guidelines.

The model recognizes four distinct states in the evidence cycle, which are each addressed in the Forensic Readiness Guidelines:

- Preparation Phase (setting up the information system)
- Running Phase (normal state)
- Awareness Phase (a problem arises, e.g. dispute)
- Investigation Phase (investigate incident)
- Learning Phase (incorporate lessons learned into system)

Forensic Readiness Guidelines define forensic readiness as "systematic advance action to prepare and install the components, systems and procedures needed to enable a company, in an efficient and cost-effective manner, to find, conserve and produce satisfactory evidence when an incident occurs."[37]

The guidelines identify 10 steps:

- Identify potential sources & different types of available evidence.
- Ensure monitoring is targeted.
- Decide which crimes and disputes, electronic evidence may be required for.
- Specify the circumstances when escalation to a full investigation is required.
- Train staff, to ensure all understand the legal consequences of incidents.
- Determine the evidence requirement.
- Establish a secure logging capability for the electronic evidence requirement.

[37] CTOSE CONSORTIUM, *CTOSE Project Results*, *op. cit.* (as in n. **??**), p. 8.

- Plan forensic procedures and adopt suitable internal/external standards.
- Set up a policy for the secure storage and handling of logs.
- Ensure that data are legally collected.

The Investigative Process Model also covers the different phases of the methodology model: preparation, running, assessment, investigation and learning phase. The process model defines the flow of actions and decisions which have to be considered or executed in the case of an IT incident.[38] An electronic version of the investigative process model comprising a flow chart model of all actions, decisions and the relationships between them was developed. This tool was called C*CAT (Cyber Crime Advisory Tool), but does not appear to be available anywhere.

Legal Advisor

Within the CTOSE project a legal advisory tool was developed with the objective of ensuring that evidence can be produced which is admissible and has probative value in court. Unfortunately, the tool does not appear to be available anymore. What follows is the published description of the legal advisory tool:

> "The legal requirements which cover the gathering of electronic evidence depend on the nature of the court or tribunal to which this evidence is to be submitted: Criminal Court – firmly set in statute and case law and specified in the Member States' penal codes, developed through time for the physical world and now thrust into the electronic. The objective of a criminal court is to determine whether the accused person has committed a crime and to punish the wrongdoer accordingly."
>
> Civil Court – set in tort, contract and property law for disputes between individuals (persons or companies) with an increasing amount of electronic evidence now involved in company disputes. Adjudication is given upon evidence presented and a remedy is provided. Tribunal or Extra-judicial – defined in industrial relations law and mediation and arbitration procedures. These instances range from tribunals which adjudicate in labour law disputes between employees and their employers to extra-judicial mediation and arbitration hearings for dispute resolution between companies.
>
> The rules for the procedure and acceptance of evidence are somewhat different for these different bodies, and for their instantiation in different countries: the elements

[38] CTOSE CONSORTIUM, *CTOSE Project Results*, *op. cit.* (as in n. **??**), p. 5.

> of the evidence and the burden of proof con vary considerably. This fact, combined with the international nature of the Internet, means that an investigator may well find himself handling evidence to be presented to a tribunal whose rules of evidence are quite unfamiliar to him. It is therefore essential to have a uniform approach to electronic evidence processing and handling. Criminal courts tend to have the most stringent requirements for evidence gathering and admissibility, because of the burden of proof required (e.g. in English law "beyond reasonable doubt", in French law "la conviction intime"). For this reason, LE bodies accustomed to producing evidence for criminal courts, and accustomed to ensuring a satisfactory chain of custody, have made a significant contribution to the project by contributing advice and "best practice".
>
> The Legal Expert Tool (Figure 3) is an on-line advisory tool for the investigator. It works interactively, asking questions and using the answers to offer the user guidance on national penal and privacy laws, to ensure that the evidence produced will be admissible, convincing, and legally gathered."[39]

Forensic Autopsy Tool (FAT)

Another tool developed within CTOSE was the FAT (Forensic Autopsy Tool). Electronic evidence has important legal aspects. In order to be presented to a court, it must be authenticated as being:[40]

- Authentic - this evidence is indeed what it is claimed to be; it has not been tampered with.
- Complete - there is nothing missing which could contribute to the understanding of the points under discussion
- Trustworthy - this evidence has been collected and handled in such a way that there can be no doubt about its authenticity and veracity.

> "The FAT was developed from an open source tool created by Brian Carrier from @stake and Purdue University. The tool addresses the problems outlined above; the team further enhanced it by developing an XML format to package an item of evidence, and display a time line report."[41]

FAT can be used for log file analysis:

[39] CTOSE CONSORTIUM, *CTOSE Project Results*, *op. cit.* (as in n. **??**), p. 4.
[40] CTOSE CONSORTIUM, *CTOSE Project Results*, *op. cit.* (as in n. **??**), p. 10.
[41] CTOSE CONSORTIUM, *CTOSE Project Results*, *op. cit.* (as in n. **??**), p. 10.

> "The FAT can be used to recover all possible logs from different sources. There is as yet no standard format for logs, nor for security reporting, and therefore it is not possible to entirely automate log file analysis. However the CTOSE project has developed a set of requirements for log file analysis tools"[42]:
>
> - The tool should recognise the event structure of the log file.
> - The tool should allow scanning / analysing of the events recorded using a configurable built-in database (which permits updating).
> - The tool should warn the user when it encounters unknown or suspicious events, and ask for a classification of these.
> - The tool should produce a statistical report covering critical, high, medium, and low level security events, including an explanation of the classification criteria.
> - The tool should allow the sorting and filtering of events as well as output of the result into a specific user-defined file format.
> - The tool should log any input-output error.
> - The tool should provide clear and secure documentation of the user's actions.

In the particular – but very common – case of analysis of log files from different sources there are the following further requirements:

- The tool should aggregate all the logs from all the different sources.
- The tool should enable mapping each log file onto its corresponding node in the network (with the network topology displayed graphically).
- The tool should allow filtering of the logs, keeping only events of a given specification.
- The tool should display a sorted time-line list of all the events gathered.

The FAT address these problems. However the challenge is significant, and although the project has developed a beta tool, further effort is required before it is fully ready for roll-out.

CTOSE Demonstrator and Project Story Board

In order to demonstrate the commercial viability of the CTOSE research results, the CTOSE demonstrator and project story board were developed.

> "The CTOSE Demonstrator shows the methodology applied in a realistic commercial setting. The demonstrator includes three scenarios, which were developed in

[42] CTOSE CONSORTIUM, *CTOSE Project Results*, *op. cit.* (as in n. **??**), p. 11.

consultation with the SIG to show the range of challenges arising from high-tech crime. The demonstrator shows these three attack scenarios being handled both with and without the CTOSE methodology."[43]

"The Story Board is a portable demonstration and sales tool, with some 250 interconnected slides and actions."[44]

Further reading

- FRINGS, S. et al., "Cyber Crime Advisory Tool - C*CAT: a holistic approach to electronic evidence processing", *Proceedings of the 10th International Conference on Human-Computer Interaction*, 3 2003
- ROWLINGSON, ROBERT, "A Ten Step Process for Forensic Readiness", *International Journal of Digital Evidence*, vol. 2 2004, Nr. 3 , `http://www.utica.edu/academic/institutes/ecii/publications/articles/A0B13342-B4E0-1F6A-156F501C49CF5`
- DINANT, JEAN-MARC, "The long way from electronic traces to electronic evidence", *International Review of Law, Computers Technology*, Vol. 18 2004
- LEROUX, OLIVIER, "Legal admissibility of electronic evidence", *International Review of Law, Computers Technology*, Vol. 18 2004, Nr. 2
- PEREZ ASINARI, MARIA VERONICA, "Legal constraints for the protection of privacy and personal data in electronic evidence handling", *International Review of Law, Computers Technology*, Vol. 18 2004, Nr. 2
- BROUCEK, VLASTI, TURNER, PAUL and FRINGS, SANDRA, *Music piracy, universities and the Australian Federal Court: Issues for forensic computing specialists*, 2005, 21 , `http://dx.doi.org/10.1016/j.clsr.2005.01.014`
- SATO, O, BROUCEK, V and TURNER, P., "Electronic evidence management for computer incident investigations: a prospect of CTOSE", *Security Manage*, 2005, Nr. 18
- BROUCEK, VLASTI and TURNER, PAUL, "Winning the Battles, Losing the War? Rethinking Methodology for Forensic Computing Research", *Journal in Computer Virology*, Vol. 2 2006, Nr. 1

[43] CTOSE CONSORTIUM, *CTOSE Project Results*, *op. cit.* (as in n. **??**), p. 11.
[44] CTOSE CONSORTIUM, *CTOSE Project Results*, *op. cit.* (as in n. **??**), p. 12.

Chapter 4

Legal Metadata

A concise introduction to metadata and it's role in electronic records management can be found in FRANKS, PAT and KUNDE, NANCY, "Why metadata matters", *The Information Management Journal*, Sept/Oct 2006.

ISO has issued a standard for record's metadata: ISO 23081 – Metadata for records. This standard is an extension of ISO 15489 (Information and documentation – Records management) and is intended to help people understand metadata from a records management perspective, not to develop a new metadata set. The standard gives a definition of records management metadata and describes its purpose. It describes roles and responsabilities in assigning and maintaining metadata.

Over the years, a large amount of metadata schemas and standards have been created. Like records themselves, metadata must be tailored to the particular business activity that generates or uses them. Considering the diversity in business activities, the numer of metadata schemas in use comes as no surprise. With the networked society, the desire to interchange information and records efficiently has given rise to more interest in ensuring interoperability of metadata.

Researching the progress in metadata standardization in general is outside the scope of this deliverable. A snapshot of important metadata standards was made by the MetaMap project.[1]

Efficiently developing standards-compliant metadata schemas for particular environments and then reusing them for other business activities was the subject of the Clever Recordkeeping Metadata Project.[2]

[1] `http://www.mapageweb.umontreal.ca/turner/meta/accueil.html`.
[2] `http://infotech.monash.edu.au/research/groups/rcrg/crkm/index.html`.

> "The aim was to bring together researchers and practitioners to investigate how standards-compliant metadata could be created once in particular application environments, then used many times to meet a range of business and recordkeeping purposes. The project wished to explore how to move away from the current resource intensive process of manual metadata attribution and stand-alone systems, towards an integrated suite of business systems and processes supporting recordkeeping functions."[3]

Though many metadata schemas exist, only few focus specifically on legal metadata. A major obstacle to overcome is of course the great disparity between legal systems amongst various countries and regions. Striving for one standardized set of legal metadata that would be valid around the globe is an unachievable goal. Interoperability of legal metadata schemas is a more realistic objective. The section on Legal Informatics discussed the state of research on those technologies that hold promise for an interoperable legal metadata framework.

In what follows, two legal domains – data protection and copyright – are discussed in more detail. Copyright is an issue that affects virtually every organisation, though some more than others. It is also one of the rare legal domains that has globally harmonized rules to such a degree. Finally, a number of metadata schemas exist that refer to copyright directly or indirectly. These schemas are briefly reviewed from the angle of compliance.

Data protection is not as highly harmonized as copyright, there are notable differences in the level of protection between countries. In the EU a comprehensive and harmonized data protection regulation exists, which has inspired a number of other countries to enact similar rules. No metadata schemes dealing with privacy and data protection have emerged in the course of this state-of-the-art review. Therefor the discussion centers around requirements for such a metadata schema, taking the EU data protection rules as the starting point.

4.1 Data protection

At present there appears to be no metadata schema in existence to capture information relevant to data protection in records management, even though much research is devoted to building privacy protection into computer systems at various levels through Privacy-Enhancing Technologies.[4] Privacy-Enhancing Technologies (PET) encompass all the technical controls that can

[3] `http://infotech.monash.edu.au/research/groups/rcrg/crkm/index.html`.

[4] See for instance the EU Commission's page on Privacy Enhancing Technologies: `http://ec.europa.eu/information_society/activities/privtech/index_en.htm`.

be used to protect personal data, including the design of the information systems architecture.[5] Much research has been devoted to privacy-friendly identity management systems[6] or languages to express privacy preferences on the web.[7]

In what follows the EU data protection rules are the focus of attention and it is from these rules that the list of questions to be addressed in records management is derived.

- Which data protection laws apply (jurisdiction)?
- Is there personal data present in the records?
- Is there 'sensitive data' present in the records?
- Do I have legitimate grounds for processing the data?
- Why am I (still) processing this data?
- How am I processing my data?
- What is my data quality assurance policy?
- What is my data security policy?
- Is my data crossing borders?

Each of these questions is discussed below. As these questions are very high level, it is quite possible that they will apply with little or no modification to data protection regulations from outside the EU. This is a subject for future research and/or validation in practice.

Notably the U.S. do not have a general data protection regime in place at present, however sector-specific privacy rules do exist.[8] For instance in the medical sector, health records are protected by the Health Insurance Portability and Accountability Act of 1996. U.S. Federal agencies are subject to the Privacy Act of 1974.[9]

Canada has separate privacy laws for the public and private sectors. The public sector is governed by the Privacy Act and Access to Information Act 1983. A federal privacy law for the private sector was introduced in 2000: The Personal Information Protection and Electronic Documents Act (PIPEDA).

[5] BORKING, JOHN, "Privacy Rules, A Steeple Chase For Systems Architects", In X. (ed.), *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, Ispra, Italy, W3C, 2006 , `http://www.w3.org/2006/07/privacy-ws/papers/`, p. 1.

[6] See for an example the Prime research project `https://www.prime-project.eu/` and the following Prime-Life project (`http://www.primelife.eu/`).

[7] See P3P, `http://www.w3.org/P3P/`.

[8] For an overview see MARCUS, J. SCOTT, CARTER, KENNETH and ROBINSON, NEIL, E.A., *Comparison of Privacy and Trust Policies in the Area of Electronic Communications*, Bad Honnef, wik-Consult GmbH, 2007 , `http://ssrn.com/abstract=1086929`, p. 57-65.

[9] This act does not afford protection to foreigners, which no doubt contributes to the European Commission's finding of an inadequate level of data protection in the U.S.

Before delving into the questions themselves a few key terms from the data protection domain are defined. Then a real world example of a multinational grappling with data protection compliance is presented, which will serve to illustrate key points further on.

4.1.1 Data protection terminology

As is the case in many legal domains, a specific terminology is used in the field of data protection. For a better understanding of what follows, a few key terms are explained.

The data subject is the person about whom data is being processed.

Processing covers the entire life cycle of data, from collection and registration through to destruction of data. The EU Data Protection Directive refers to collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction.[10]

The data controller is the person or organization in charge of the processing of data. The criterium used in the EU to identify the data controller is who determines, alone or jointly with others, the purposes and means of the processing of personal data.[11] Anyone who works on behalf of the data controller is a processor.

4.1.2 Learning from the experience of others: the SWIFT case

SWIFT is an industry owned cooperative supplying secure, standardised messaging services and interface software to over 7,800 financial institutions worldwide. SWIFT is solely a messaging intermediary for transmitting secure and confidential financial messages between financial institutions. SWIFT is not a bank, nor does it hold accounts of any customers.[12]

SWIFT is a company based in Belgian with subsidiaries in many countries, including Australia, Brazil, Switzerland, Germany, Spain, France, the United Kingdom, Hong Kong, Ireland, Italy, Japan, Luxembourg, Sweden, Singapore, South Africa, and the United States. SWIFT has two data centres, one located in the EU and the other in the United States. All data is mirrored on both locations as a back-up measure.[13]

[10] Art. 2 b EU Data Protection Directive.
[11] Art. 2 d EU Data Protection Directive.
[12] Swift Press Release, "SWIFT statement on compliance policy", June 23d 2006, `http://www.swift.com/index.cfm?item_id=59897`
[13] Article 29 Working Party Opinion 10/2006, p. 8.

An article appeared in the June 23 edition of New York Times, the Wall Street Journal and other U.S. journals on terrorism investigations and the role of SWIFT, sparking media debate on the issue of data protection and data transfer in Europe and the U.S.

Personal data, collected and processed via the SWIFT network for international money transfers had been provided to the United States Department of the Treasury ("UST") since the end of 2001 in response to compulsary administrative subpoenas under American law for terrorism investigation purposes. Under US law, an administrative subpoena is an order from a government official to a third party, instructing the recipient to produce certain information.[14] The subpoenas addressed to SWIFT were very wide in scope, materially, territorially and in period of time under scrutiny.[15]

SWIFT negotiated with the UST on how it would organise its compliance with the subpoenas, and claims to have received "significant protections and assurances as to the purpose, confidentiality, oversight and control of the limited sets of data produced under the subpoenas."[16]

The messages of interest to the UST contain names of the beneficiary of a bank transfer, the ordering customer and information about the transfer itself, in a structured or unstructured form.[17] Clearly this constitutes personal data within the meaning of Directive 95/46/EC on Data Protection. When the processing of personal data is carried out by an organization established on the territory of an EU Member State, its data protection rules apply. [18]

> "The critical decisions on the processing of personal data and transfer of data to the UST were decided by the head office in Belgium. As a consequence, the processing of personal data by SWIFT is subject to Belgian law, implementing the Directive, regardless of where the data processing takes place."[19]

Given SWIFT's level of autonomy in dealing with the personal data in its possession, SWIFT must be regarded as a 'data controllor' upon whom all the data protection obligations are incumband, and not merely a data processor operating under supervision and responsibility of another entity.[20] The Belgian Data Protection Authority ruled on the case at hand on September 27th 2006.

[14] Article 29 Working Party Opinion 10/2006, p. 8.
[15] Article 29 Working Party Opinion 10/2006, p. 8.
[16] Swift Press Release, "SWIFT statement on compliance policy", June 23d 2006, `http://www.swift.com/index.cfm?item_id=59897`
[17] Article 29 Working Party Opinion 10/2006, p. 8.
[18] Art. 4 §1 a of the EU Data Protection Directive.
[19] Article 29 Working Party Opinion 10/2006, p. 9.
[20] Article 29 Working Party Opinion 10/2006, p. 10.

Given the high profile nature of the case, the European Commission decided to follow the case as well in collaboration with the Member States. Notably, the question arose whether banks affiliated with SWIFT are in compliance with their national laws on data protection when they use the system for the processing of payments.

The inquiries of the various EU data protection agencies were coordinated in the Article 29 Working Party, which resulted in the publication of Opinion 10/2006.

The Article 29 Working Party states that "By deciding to mirror all data processing activities in an operating centre in the US, SWIFT placed itself in a foreseeable situation where it is subject to subpoenas under US law." The fact that compliance with the US subpoenas is mandatory and lawful in the US does not excuse SWIFT from its obligations under Belgian law.

Though the Data Protection Directive and the Belgian Data protection Act state that data processing has legitimate grounds when it is imposed by a legal obligation to which the controller is subject, this does not cover foreign rules from outside the EU.[21]

The European Commission has issued clarification on safeguards to be put in place by a recipient of personal data in a third country so that transfer may occur even if the level of protection offered by legislation is insufficient. The documents address transfers from a data controller to a data controller, from a data controller to a processor[22] and transfers within multinationals.[23]

4.1.3 Jurisdiction: which data protection laws apply?

Organizations operating strictly on the territory of one country must comply with the data protection rules of that country, and in principle of that country alone. Organizations, established in a single country, but operating across national borders, may have to take into account data protection legislation from different countries. Multinationals by definition must comply with the laws of more than one country.

When it comes to the territorial scope of laws, there is no uniform way to determine which country's law apply to the exclusion of all others. In the end, a country can more or less arbitrarily determine the territorial reach of its laws (enforcement being a different matter). In practice, there must usually some connection to a country before its laws will apply, e.g. activity on the country's territory or by a country national, . . .)

[21] Article 29 Working Party Opinion 10/2006, p. 18.

[22] See `http://ec.europa.eu/justice_home/fsj/privacy/modelcontracts/index_en.htm`.

[23] Article 29 Working party Working document WP 74, "Transfers of personal data to third countries: Applying Article 26(2) of the EU Data Protection Directive to Binding Corporate Rules for International Data Transfers" adopted by the Working Party on 3d June 2003 and further complementary documents WP 107 and WP 108.

The reach of the EU data protection directive does not depend on whether the data subject is a EU citizen or not. Nor is it decisive whether the records reside in the EU or whether the processing occurs in the EU. The reach of the EU data protection directive depends **primarily** on the fact whether the data controller is established in an EU Member State. Each Member State shall apply its national data protection provisions when data processing is carried out in the context of the activities of an establishment of the controller on its territory.[24] **Secondarily**, the Member States must apply data protection rules to any processing of data occuring on their territory, even if the data controller is not established in the EU. [25] Knowing which jurisdictions apply will also tell which data protection authorities have oversight authority over an organization's data processing activities as well as any requirements to notify data protection authorities of the fact that data processing is planned.

In the SWIFT case, the contested processing of personal data (handing over personal data to the UST) occurred in the U.S. However, the Belgian Data Protection Act is deemed applicable on SWIFTs decisions regarding data processing, wherever the actual processing occurs, because SWIFT is established in Belgium.

Companies established in the U.S. and processing data in the U.S. usually have nothing to fear from EU data protection regulation. However, if a U.S. company starts collecting and processing personal data in the EU, it must take into account EU data protection rules.

4.1.4 Is there personal data present in the records?

Ascertaining wether or not records contain personal data is not an easy task, furthermore, different countries have different legal definition(s) of what personal data is.

In theory, the definition is harmonised amongst all the EU Member States, in practice the interpretation differs.[26] There is little debate where data about identified persons is concerned, this is obviously personal data. Where to draw the line between indirectly identifiable personal data and truly anonymous or non-personal data is a source of discord. Root of the discussion is the effort required to identify the data in question. Recital 26 of the EU Data Protection Directive states that "account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person". Belgium has taken this to mean

[24] By extension, when a Member States law is applied outside its territory, by virtue of the rules of international public law, e.g. on ships or airplanes operating under a country's flag, the data protection rules must be applied there as well. Art. 4 EU Data Protection Directive.

[25] Art. 4 EU Data Protection Directive.

[26] Article 29 Working Party, Opinion 4/2007 on the concept of personal data, June 2007.

that as long as someone out there can identify the data, it is personal data to which the law applies, no matter how unlikely it is that this person would cooperate with the data controller. A similar position is held by the Swedish and French data protection authorities. Other countries, like Germany, focus on the effort it would cost the data controller himself to identify the data, including the cost or difficulties in enlisting the help of others.

4.1.5 Is there 'sensitive data' present in the records?

The EU Data Protection directive affords more stringent protection to "sensitive data", which is data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership and data concerning health or sex life. The processing of the sensitive data is only allowed on grounds explicitly mentioned in Art. 8 (2)-(7) of the EU Data Protection Directive. Again, the Member States may differ slightly in their implementation of this more stringent protection regime for personal data.

Most of the data processed by SWIFT would not qualify as sensitive data, though exceptions are possible. Payment of a membership fee to a trade union could be considered 'sensitive data' under the EU Data Protection rules, especially if the included free text form explicitly states it is the payment of a membership fee. No doubt examples for the other types of sensitive data exist as well.

An organisation will need to consider at which level of granularity it can safely make the assessment of whether sensitive data is involved. Is it necessary to tag individual records as containing sensitive data or can a tag be added at series or collection level? Is there a difference from one country to the other?

4.1.6 Do I know about who I'm processing personal data

One might assume that if an organisation processes personal data, it must know exactly about who. However, this is not necessarily so since the legal definition of personal data covers all data that **can** be tied to a person, whether or not this person is already identified or only potentially identifiable. An analysis of the level of data subject identification in the organisation can clarify to which degree data subjects are identified or identifiable.[27]

[27] With regard to database design, see GOUNARIS, ANASTASIOS and THEODOULIDIS, BABIS, "Data Base Management Systems (DBMSs): Meeting the requirements of the EU data protection legislation", *International Journal of Information Management*, 23 2003, Nr. 3 , http://dx.doi.org/10.1016/S0268-4012(03)00023-9, p. 189.

The personal data processed by SWIFT in its financial transfer messages is almost exclusively data about identified persons. Bank account numbers are generally linked to identified persons, with the exception perhaps of bank account numbers owned by legal persons. Possibly, free text portions of financial transfer messages may contain personal data about persons which are not identified but may be identifiable.

A secondary question is whether or not I am able to tie all dispersed personal data I have about one person together if required. The EU Data Protection Directive requires a data controller to be able to inform each data subject as to whether or not data relating to him are being processed as well as communication in an intelligible form of the data being processed (right of access).[28] The report given in response to an access request must contain additional information about why and how data is being processed, as will be discussed further on.

4.1.7 Do I have legitimate grounds for processing the data?

The EU Data Protection directive prohibits processing of personal data unless the controller has legitimate grounds to do so. Legitimate grounds are one of the following:

"Art. 7 Member States shall provide that personal data may be processed only if:

- (a) the data subject has unambiguously given his consent; or
- (b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract; or
- (c) processing is necessary for compliance with a legal obligation to which the controller is subject; or
- (d) processing is necessary in order to protect the vital interests of the data subject; or
- (e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller or in a third party to whom the data are disclosed; or

(f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection under Art. 1 (1)."[29]

[28] Art. 12 EU Data Protection Directive.
[29] Art. 7 EU Data Protection Directive.

It should be noted that particular instances of data processing may fall under different legitimacy grounds. This is important as the grounds for legitimacy which applies also determines the limits of permissible data processing.

In the SWIFT case, the Belgian DPA made a distinction between the regular operation of its service and the transfer of data to the UST in response to subpoenas. Processing bank transfer statements exchanged between member banks was deemed legitimate because it was done in the execution of a contract with the member banks (who in turn have a contractual relation with the data subject).[30]

With respect to the transfer of data to the UST, the Belgian DPA rejected the grounds of compliance with a legal obligation by considering that foreign rules from outside the EU are not covered.[31] This position is in line with a previous opinion of the Article 29 Working party and was later confirmed by the subsequent opinion on the SWIFT case.[32] Both the Belgian DPA and Article 29 Working party did accept the grounds of legitimate interests pursued by SWIFT[33] , however both found SWIFT to be in breach of data protection regulation due to the circumstances of how the transfer to the UST was handled.

If my grounds for legitimacy is 'unambiguous consent'[34] do I want metadata that points to the actual consent given by a data subject or only that points to the policy for obtaining such consent? The same data subject may have to give consent several times, for instance because separate consent may be required for processing sensitive data or for transferring data to a third country with an inadequate level of data protection rules.[35] How do I want to deal with this in my metadata schema?

If my grounds for legitimacy is '(pre-)contractual relationship' do I want to point to the actual contract or negotions documentation with the data subject or only to the business process from which the contract arises?

In case my grounds for legitmacy is 'legal obligation' do I point directly to the law(s) in question or to a company policy document interpreting and applying the law, or both?

If 'vital interests' of the data subject are the grounds for legitimacy, should metadata point to a document stating which vital interests for the individual involved or is a reference to a policy

[30] Belgian DPA, Decision 37 of 2006, p. 16.
[31] Belgian DPA, Decision 37 of 2006, p. 19.
[32] See Article 29 Working Party Opinions 1/2006, p. 8 and 10/2006, p. 18
[33] Belgian DPA, Decision 37 of 2006, p. 20 and Article 29 Working Party Opinion 10/2006, p. 18.
[34] Regarding the conditions of obtaining consent, notably by automated processes, see BORKING, *Privacy Rules, A Steeple Chase For Systems Architects*, *op. cit.* (as in n. **??**), p. 10 ff.
[35] GOUNARIS and THEODOULIDIS, *International Journal of Information Management 23 [2003]*, *op. cit.* (as in n. **??**), p. 189.

document sufficient?

If 'task of public interest or exercise of official authority' is my grounds for legitimacy to what level of detail should this be reflected in metadata?

If 'legitimate interests' are the grounds for legitimacy, how can this be captured in metadata? Additionally, to what extent should metadata point to information explaining why in the business process at hand the legitimate interests of the data controller are not overridden by the interests of the data subjects?

The banks making use of the SWIFT service to transfer messages in relation to financial transfers between financial institutions have a contractual relation with the client putting the order for the financial transfer. Where the recipient of the financial transfer is concerned, the banks have a legitimate interest in processing his data at least for the completion of the transfer. Subsequent processing of the recipients data by the bank, for instance preservation for a number of years may be required by law, e.g. accounting laws or anti-money laundering laws.

SWIFT has no direct contractual relation with either the client initiating the financial transfer or the recipient, thus it could at best claim an indirect contractual relationship with the data subjects via the member banks or rely on 'legitmate interests'.

Finally, there is a temporal aspect to consider. In particular where consent is the grounds for legitimacy, consent may be withdrawn at a certain point by the data subject. This affects the kinds of processing which are still allowed or even required after this occurs. The other grounds for legitimacy may also be valid only for a specific period in time. A legal obligation to preserve records is grounds to process personal data, but only until the preservation term runs out. Determining how long the data controller can claim his legitimate interests prevail upon those of the data subject is a delicate issue.

4.1.8 Why am I (still) processing this data?

All the obligations incumbant upon the data controller hinge upon the purpose for which he is processing personal data.

Personal data may only be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes.[36] Personal data should be adequate, relevant and not excessive in relation to the purpose of processing.[37] Being able to make the connection between (categories of) personal data and the reason for which it is being processed in

[36] Art. 6 §1 b EU Data Protection Directive.
[37] Art. 6 §1 c EU Data Protection Directive.

metadata can help ensure that no unnecessary, irrelevant or inadequate data is being processed. If no reason for the data being there is stated, at least the question should be raised as to why it is there in the first place. Obviously, merely referring to a general or vague purpose will offer no guarantees of actual compliance with data protection. Ideally, there should be a link with one or more specific business processes or goals which justify which justify processing of that particular set of personal data. In case a data subjects requests access to their personal data, the data controller should be able to explain in the report what the purpose of the processing is as well as the categories of data concerned.[38]

The simplest data processing scenario is one where data is collected from the data subject, processed and then discarded. An example is an organization that holds a survey amongst customers, compiles the results into a report and then discards or anonymizes the source data. In reality, this scenario is rather rare. Usually personal data is obtained either from the data subject or from another source for a specific purpose and preserved. At a later time, the data controller may realise that the same data can be used for other purposes as well and wish to do so. Such reuse or 'further processing' is permissible under the EU Data Protection rules only if the new purpose is compatible with the original purpose for which the data was obtained. Note that reuse of data for historical, statistical or scientific purposes shall not be deemed incompatible, inasfar as the conditions imposed by the Member States upon such reuse are followed.[39]

Reuse of personal data for a new and incompatible purpose is not illegal per se, but it must conducted as if one were starting from scratch. A ground for legitimacy must be found (e.g. consent for the new purpose), the data subject must be informed about the data processing, the Data Protection Authority may need to be notified, etc.

SWIFTs original purpose for processing personal data is to route financial transfer messages between member banks, a purely commercial purpose. Handing over the data to the UST in response to subpoenas for alleged terrorism investigations is an example of further processing which is incompatible with the original purposes.[40] The transfer of data to the UST could have been lawful, if SWIFT had complied with the relevant provisions of the Belgian Data Protection Act as if the data in question were freshly collected.[41]

Personal data may be preserved in identifiable form for only as long as it is necessary for the pur-

[38] Art. 12 EU Data Protection Directive.

[39] Art. 6 §1 b) EU Data Protection Directive. For an overview of the implementation of this provision by a selection of EU Member States, see IACOVINO, LIVIA and TODD, MALCOLM, "The long-term preservation of identifiable personal data: a comparative archival perspective on privacy regulatory models in the European Union, Australia, Canada and the United States", *Archival Science*, vol. 7 2007, Nr. 1 , `http://dx.doi.org/10.1007/s10502-007-9055-5`, p. 112 ff.

[40] See Article 29 Working Party Opinion 10/2006 p. 15.

[41] See Article 29 Working Party Opinion 10/2006, p. 15-16.

pose for which it was collected and further processed. After this time, it may only be preserved in anonymized form.[42]

The EU Data Protection Directive gives individuals the means to enforce this principle by granting each data subject the right to demand that data held in breach of the data protection rules is erased or blocked from the system.[43] Furthermore, any third parties to whom the data has been disclosed must be notified of the erasure or blocking.[44]

Not only may the data subject demand that personal data processed in breach of the rules is removed, he may even demand that data processed in compliance is no longer used. This is called the data subject's right to object.[45] The objection must be justified based upon compelling legitimate grounds relating to the data subject's particular situation. The Member States may impose further limits on the right to object. The Belgian DPAct does not allow for a right to object when the data processing is required in a (pre-)contractual relationship with the data subject or in light of a legal obligation.[46]

Note that in case the anticipated purpose of the data processing is direct marketing the data subject may object at will, without the need for any justification.[47]

Without going into the details of when a data subject may object to his data being processed any longer. It is important to consider the impact of such an objection upon records management. Simply deleting information may not be enough, as this alone would not prevent the same person's data to be collected and processed all over again at a later time. Multinational organizations will need to figure out whether the objection is valid for all data processing activities in the world or region, or whether it only applies to local data processing.

In the SWIFT case, the financial message transfer service was deemed to be regulated by the Belgian Data Protection Act. As a result any EU citizen would presumably have the right to object to his data by processed for this purpose by SWIFT in accordance with the Belgian Data Protection Act, and not his home country's act. By contrast, the various sales offices located in the EU countries would probably be subjected to local data protection rules for the processing of data about their respective employees.

The data controller should ask himself:

- Do I need this data? Why?

[42] Art. 6 §1 §1 e) EU Data Protection Directive.
[43] Art. 12 b) EU Data Protection Directive.
[44] Art. 12 c) EU Data Protection Directive.
[45] Art. 14 a EU Data Protection Directive.
[46] Art. 12 §2 BDPAct.
[47] Art. 14 b) EU Data Protection Directive.

- Do I need all of it?
- Is it sufficient for my needs?
- For how long do I need it?

Ideally, the metadata associated with personal data would quickly lead to the policy documents that answer these questions.

4.1.9 What is my data quality assurance policy?

The data controller shall take steps to ensure that personal data is accurate and, where necessary, kept up to date. Every reasonable step must be taken to ensure that data which are inaccurate or incomplete are erased or rectified, taking in to account the purposes for which they were collected or for which they are further processed.[48]

The data subject may demand that inaccurate information held about him is either rectified, erased or blocked by the data controller.[49] Additionally, any third parties to whom the data has been disclosed must be notified of the rectification, erasure or blocking unless this proves impossible or involves a disproportionate effort..[50]

4.1.10 What is my data security policy?

The confidentiality and security of personal data must be protected with all reasonable means.[51] Not only must access by third parties be strictly controlled, access by employees of an organization should be limited to a need-to-know basis.[52] Granting access could be determined based on the role someone plays with respect to the organization. If this is the case, metadata accompanying personal data should clarify which roles have access to which data at which times.

It should be noted that the data subject is granted a right of access to all personal data processed about him by the data controller, as indicated above.[53] The report given in response to an access request must explain which employees or categories of employees have access to the data.[54] Oftentimes, personal data of several people is tied together somehow. Granting one data subject

[48] Art. 12 d) EU Data Protection Directive.
[49] Art. 12 b) EU Data Protection Directive.
[50] Art. 12 c) EU Data Protection Directive.
[51] Art. 17 EU Data Protection Directive.
[52] With regard to database design, see GOUNARIS and THEODOULIDIS, *International Journal of Information Management 23 [2003]*, *op. cit.* (as in n. **??**), p. 187.
[53] Art. 12 a) EU Data Protection Directive.
[54] GOUNARIS and THEODOULIDIS, *International Journal of Information Management 23 [2003]*, *op. cit.* (as in n. **??**), p. 190.

access to his personal data must be done without unlawfully disclosing personal data about others.

Implementing audit trail functionality in the records management application appears to be a minimal requirement in order to be in compliance with data protection regulation, not only to ensure the confidentiality of the data but also to enforce many of the other obligations incumbent upon the data controller. An audit trail aims to keep track of who accessed, amended, deleted or disseminated data, when and for what purpose.[55]

4.1.11 How am I processing my data?

At issue here are the relevant circumstances of data processing. Am I processing data in compliance with all relevant legal obligations?

What metadata do I need (or want to invest in having) to quickly reflect that processing is compliant?

The EU Data Protection Directive imposes a number of obligations to ensure transparancy of data processing towards the data subject concerned, notably by obliging the data controller to provide information about the processing of his data. At least information pertaining to the identity of the data controller and where applicable his representative, the intended purposes of processing, the (categories of) recipients of data as well as any additional information required to ensure fair processing considering the specific circumstances. Also, the data subject must be informed that he has a right of access and rectification of his personal data. In case the data subject already has all this information at his disposal, the obligation to inform is waived. A distinction is made between cases where data is collected directly from the data subject and cases where it is obtained from other sources. In the latter case the data controller may sometimes delay informing the data subject until such time as he discloses the data to a third party. Also, the data controller must disclose categories of data he has obtained about the data subject from his source. Finally, when the data controller collects personal data from a source other than the data subject himself, the obligation to inform is waived when the provision of such information proves impossible, would involve a disproportionate effort or if recording or disclosure is expressly laid down by law.[56] Without going into the nuances of exactly when to inform the

[55] With regard to database design, see GOUNARIS and THEODOULIDIS, *International Journal of Information Management 23 [2003]*, *op. cit.* (as in n. **??**), p. 191. See also RUNDLE, MARY, "International Personal Data Protections and Digital Identity Management Tools", In X. (ed.), *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, W3C, 2006 , `http://www.w3.org/2006/07/privacy-ws/papers/21-rundle-data-protection-and-idm-tools`, p. 3.

[56] Art. 10 - 11 EU Data Protection Directive.

data subject and what to inform him about according to the data protection rules of de different Member States, it is clear that metadata capturing what initiatives the organization has in fact taken to inform data subjects is of great value. An open question is whether I want or need metadata that points to the actual information given to a data subject or only that points to the policy for providing such information? Probably, in some business processes the first solution will be worth the effort (e.g. processing employee records), while in others not (e.g. collecting customer feedback via a web form).

Another provision intended to foster transparancy of personal data processing is the obligation to notify the competent data protection authority before starting the processing of data.[57] The notifications are to be kept in a public register.[58] The Directive affords the Member States plenty of freedom to determine when notification is or is not required, as a result considerable differences between the Member States may exist.

Use of automated decision making processes based on personal data is restricted by the EU Data Protection Directive[59] . When such automated decision processes are employed, the data subject may request an explanation of the logic behind it.[60]

4.1.12 Do I know where data comes from and where it goes?

In case a data subjects requests access to their personal data, the data controller should be able to give information as to their source (if available) as well as report who the recipients or categories of recipients are to whom the data are disclosed.[61]

Any third parties to whom personal data has been disclosed must be notified of rectification, erasure or blocking of such data unless this proves impossible or involves a disproportionate effort.[62]

4.1.13 Is my data crossing borders?

This question is distinct from the issue of jurisdiction, though of course both issues are related. The EU Data Protection Directive, and as a consuence all implementing legislations, restrict the transfer of personal data to countries that do not fall with the reach of the Directive. In principle,

[57] Art. 18 EU Data Protection Directive.
[58] Art. 21 EU Data Protection Directive.
[59] Art. 15 (1) EU Data Protection Directive.
[60] Art. 12 EU Data Protection Directive.
[61] Art. 12 EU Data Protection Directive.
[62] Art. 12 c) EU Data Protection Directive.

transfer to such countries is only permissible if it provides an adequate level of protection to personal data.[63] This is to prevent easy circumvention of the obligations by 'off-shoring' data processing operations.

By way of exception and under the conditions laid down by the Member States, personal data may be transfered to third countries which do not provide a sufficient level of protection if:

"(a) the data subject has given his consent unambiguously to the proposed transfer; or

(b) the transfer is necessary for the performance of a contract between the data subject and the controller or the implementation of precontractual measures taken in response to the data subject's request; or

(c) the transfer is necessary for the conclusion or performance of a contract concluded in the interest of the data subject between the controller and a third party; or

(d) the transfer is necessary or legally required on important public interest grounds, or for the establishment, exercise or defence of legal claims; or

(e) the transfer is necessary in order to protect the vital interests of the data subject; or

(f) the transfer is made from a register which according to laws or regulations is intended to provide information to the public and which is open to consultation either by the public in general or by any person who can demonstrate legitimate interest, to the extent that the conditions laid down in law for consultation are fulfilled in the particular case."[64]

Another option is for the data controller to construct an adequate level of protection for data being transferred[65] , for instance through contractual provisions or binding corporate rules.[66] In this case, transfers must be authorized by the data protection authority and notified the European Commission. The European Commission has issued three decisions on standard contractual clauses that construct adequate safeguards, two of which regulate transfers from a data controller to a data controller while the third regulates transfers from a data controller to a processor.[67]

By mirroring its data centre in the U.S. SWIFT effectively exported all of its personal data to a country lacking an adequate level of protection. No steps were taken by SWIFT to construct an adequate level of protection through contractual provisions or binding corporate rules, nor does any of the exceptions apply.[68]

[63] Art. 25 EU Data Protection Directive.
[64] Art. 26 §1 EU Data Protection Directive.
[65] Art. 26 §2 EU Data Protection Directive.
[66] See Article 29 Working Party Opinion 10/2006, p. 22.
[67] See Article 29 Working Party Opinion 10/2006, p. 22.
[68] See Article 29 Working Party Opinion 10/2006, p. 21 ff.

4.1.14 Translation of privacy rules into information systems design

The first step in building a privacy compliant information system is a privacy threat analysis. The threats identified have to be countered and neutralized in the design of the information system.[69] Privacy law rules must be incorporated into the system design, one way to achieve this is by representing the data protection rights and duties into the system.[70] Borking proposes breaking down legal texts into the smallest constituant parts (subject-verb-object) – whilst maintaining a link with their legal source – and formalizing these into a privacy ontology.[71] This approach was used in the PISA project.[72] The prototype developed in the PISA project appears to focus on a situation where lawfullness of processing depends entirely upon the consent of the data subject.[73] Further development would then be required to incorporate other situations in the system.

One aspect of designing privacy compliant systems is implementing privacy aware access control policies, their management and enforcement.[74] Two platform-independant privacy policy languages for access control are XACML and EPAL.[75]

XACML (Extensible Access Control Markup Language) is an OASIS XML standard comprising a policy language describing general access control requirements and a corresponding request/response language, which allows to query whether a particular action is permitted.[76] An extension to XACML exists with which to express the purpose for which the data was collected

[69] BORKING, *Privacy Rules, A Steeple Chase For Systems Architects*, *op. cit.* (as in n. **??**), p. 4, citing BORKING, J.J. et al., *Methodology of Privacy Threat Analysis*, The Hague, EU PISA project IST-2000-26038, 2001, Deliverable 7 of WP 2

[70] BORKING, *Privacy Rules, A Steeple Chase For Systems Architects*, *op. cit.* (as in n. **??**), p. 11.

[71] BORKING, *Privacy Rules, A Steeple Chase For Systems Architects*, *op. cit.* (as in n. **??**), p. 11.

[72] BORKING, *Privacy Rules, A Steeple Chase For Systems Architects*, *op. cit.* (as in n. **??**), p. 11, citing KENNY, S. and BORKING, J., "The Value of Privacy Engineering", *The Journal of Information, Law and Technology*, 2002, Nr. 1 , `http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2002_1/kenny/`.

[73] See BORKING, *Privacy Rules, A Steeple Chase For Systems Architects*, *op. cit.* (as in n. **??**), p. 13 ff.

[74] CASASSA MONT, MARCO; X. (ed.), *On the Need to Explicitly Manage Privacy Obligation Policies as Part of Good Data Handling Practices*, Ispra, Italy, W3C, 2006 , `http://www.w3.org/2006/07/privacy-ws/papers/`, p. 2.

[75] For an comparison of these languages, see ANDERSON, ANNE, *A Comparison of Two Privacy Policy Languages:EPAL and XACML*, Sun Microsystems Laboratories, 2005 , `http://research.sun.com/techrep/2005/smli_tr-2005-147/TRCompareEPALandXACML.html`.

[76] MADSEN, PAUL, CASASSA MONT, MARCO and WILTON, ROBIN, "A Privacy Policy Framework - A Position paper for the W3C Workshop of Privacy Policy Negotiation", In X. (ed.), *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, Ispra, Italy, W3C, 2006 , `http://www.w3.org/2006/07/privacy-ws/papers/`, p. 2 ff. See also `http://xml.coverpages.org/xacml.html`.

and the purpose for which access is requested.[77]

EPAL (The Enterprise Privacy Authorization Language) is an interoperability language for exchanging privacy policies in a structured format between applications or enterprises.[78]

Before such formal privacy policy languages can be used, formal models must be derived from natural language texts of regulations. In essence, from the full legal text portions relevant to access control are selected and translated into formal expressions.[79]

> "The Open Digital Rights Language (ODRL) is a "vocabulary for the expression of terms and conditions over digital content including permissions, constraints, obligations, conditions,offers and agreements with rights holders." The ODRL specification supports an extensible language and vocabulary (data dictionary) for the expression of terms and conditions over any content including permissions, constraints, requirements, conditions, and offers and agreements with rights holders."[80]

Besides the issue of access control, systems should support compliance with other data protection obligations. 'Privacy obligations' are policies dictating constraints, duties and expectations to data recipients, expressing how personal data should be handled. This includes amongst others data retention management, deletion of data, notifications, data transformations.[81] Privacy obligations stem from data protection rules primarily, but may also stem from the privacy preferences expressed by the data subject when giving consent.

---

[77] MADSEN, CASASSA MONT and WILTON, *A Privacy Policy Framework - A Position paper for the W3C Workshop of Privacy Policy Negotiation*, *op. cit.* (as in n. **??**), p. 3. "Privacy policy profile of XACML v2.0", available at `http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-privacy_profile-spec-os.pdf`.

[78] MADSEN, CASASSA MONT and WILTON, *A Privacy Policy Framework - A Position paper for the W3C Workshop of Privacy Policy Negotiation*, *op. cit.* (as in n. **??**), p. 3 ff.

[79] GUNTER, CARL A., "Ensuring Privacy Conformance in Inter-Domain Systems", In X. (ed.), *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, Ispra, Italy, W3C, 2006 , `http://www.w3.org/2006/07/privacy-ws/papers/`, p. 1 ff. and MAY, MICHAEL J., GUNTER, CARL A. and INSUP, LEE, "Privacy APIs: Access Control Techniques to Analyze and Verify Legal Privacy Policies", In X. (ed.), *Computer Security Foundations Workshop*, Venice, Italy, 2006 , `http://seclab.uiuc.edu/pubs/MayGL06.pdf`.

[80] MADSEN, CASASSA MONT and WILTON, *A Privacy Policy Framework - A Position paper for the W3C Workshop of Privacy Policy Negotiation*, *op. cit.* (as in n. **??**), p. 4.

[81] CASASSA MONT, *On the Need to Explicitly Manage Privacy Obligation Policies as Part of Good Data Handling Practices*, *op. cit.* (as in n. **??**), p. 2, referring to CASASSA MONT, MARCO, "Dealing with Privacy Obligations: Important Aspects and Technical Approaches", In KATSIKAS, SOKRATIS K., LOPEZ JAVIER PERNUL GÜNTHER (ed.), *Trust and Privacy in Digital Business*, Volume 3184, Lecture Notes in Computer Science, Springer, 2004 , `http://dx.doi.org/10.1007/b99832` and CASASSA MONT, MARCO, *A System to Handle Privacy Obligations in Enterprsies*, HP, 2005, HP Technical Report , `http://www.hpl.hp.com/techreports/2005/HPL-2005-180.html`.

> "A proper language is required to describe a broad variety of "events" (beyond access control, i.e., time-based events, context-based events, etc.) that might trigger obligations: this language must also allow for an explicit description of the "target" of an obligation (i.e., the personal data that is subject to the obligations) along with the actions to be carried on (e.g., deletion, notification, etc.), allowed exceptions and be extensible to future needs."[82]

Casassa Mont draws up a list of requirements and proposes further steps for the development of such a privacy obligations language.[83] The work done in the Prime project is referred to as a starting point.[84]

Building in compliance requires more than just determining what ought to be done in policies. The next logical step is to ensure enforcement of these policies. A major element in enforcement is auditability of information handling, which should result in accountability.[85]

Further reading

- BORKING, J.J., "The status of Privacy Enhancing Technologies", In NARDELLI, E., POSADZIEJEWSKI, S. and TALAMO, M. (ed.), *Certification and Security in E-Services, From E-Government to E-Business*, Boston, Kluwer, 2003
- BLARKOM, G.W. VAN, BORKING, J.J. and OLK, J.G.E., *Handbook of Privacy and Privacy-Enhancing Technologies, The case of Intelligent Software Agents*, The Hague, College bescherming persoonsgegevens, 2003 , `http://www.andrewpatrick.ca/pisa/handbook/Handbook_Privacy_and_PET_final.pdf`
- CASASSA MONT, MARCO, "Dealing with Privacy Obligations: Important Aspects and Technical Approaches", In KATSIKAS, SOKRATIS K., LOPEZ JAVIER PERNUL GÜNTHER (ed.), *Trust and Privacy in Digital Business*, Volume 3184, Lecture Notes in Computer Science, Springer, 2004 , `http://dx.doi.org/10.1007/b99832`, *op. cit.* (as in n. **??**)
- CASASSA MONT, MARCO, *A System to Handle Privacy Obligations in Enterprsies*, HP, 2005, HP Technical Report , `http://www.hpl.hp.com/techreports/2005/HPL-2005-180.html`, *op. cit.* (as in n. **??**)

[82] CASASSA MONT, *On the Need to Explicitly Manage Privacy Obligation Policies as Part of Good Data Handling Practices*, *op. cit.* (as in n. **??**), p. 2.

[83] CASASSA MONT, *On the Need to Explicitly Manage Privacy Obligation Policies as Part of Good Data Handling Practices*, *op. cit.* (as in n. **??**), p. 3 ff.

[84] PRIME Project: Privacy and Identity Management for Europe, European RTD Integrated Project under the FP6/IST Programme, `http://www.prime-project.eu/`, 2006.

[85] RUNDLE, *International Personal Data Protections and Digital Identity Management Tools*, *op. cit.* (as in n. **??**), p. 3; WEITZNER et al., *Transparency and End-to-End Accountability: Requirements for Web Privacy Policy Languages* (as in n. **??**).

4.2 Copyright

Copyright is all but inescapable for any organisation, though some sectors are more affected than others. For content producing and distributing companies copyright lies at the heart of their business. With the shift from paper to electronic records, copyright law has greatly increased it's influence on the activities of organisations outside the media sector. For instance DNV receives ships plans in the course of its business, which are copyrighted. Paper records can be read and archived without copying them, whilst mere display of an electronic record requires copying, albeit temporaliy, thus triggering copyright law.

The application of copyright law poses a number of legal questions in relation to preservation and (re)use of these documents. For the sake of clarity, these questions are addressed in relation to the main archiving processes: ingest, preservation and dissemination. Use of information is not an archival process but it is an integral part of the life-cycle of documents, therefor it will be addressed briefly.

4.2.1 Ingest

Copyright law – as the word suggest – governs the making of copies. Ingest of electronic files entails creation and preservation of at least one copy, thus triggering the application of copyright rules. The EU Copyright Directive (2001/29/EU) regulates any and all copies made of protected works: "direct or indirect, temporary or permanent reproduction by any means and in any form" (Art. 2). For example, permission from the copyrightholder(s) is required to store the plans of a ship in DNV's archive.

It should be noted that the copyright laws of all countries contain exceptions which allow copying without obtaining the copyrightholder's permission. Generally, these exceptions only apply outside any economical activity.[86]

In the interest of compliance, copyright-relevant information should be recorded at the time of ingest into the corporate archive.

- Who created the document?
- Why am I receiving it into my corporate archive?
- When did I receive it?
- Where did I receive it?

[86] Exceptions benefitting the press allow for use of copyrighted material without obtaining permission, regardless of the fact that this may be in the course of a for-profit activity. See for instance Art. 5 §3 c EU Copyright Directive.

- Who is transferring the document to me and why?
- Who claims ownership of the document?
- Do I claim applicability of a copyright exception? Which one, which jurisdiction? What is the scope of the exception: store, access by specific persons, access internally in the company, distribute externally, reuse without external distribution, reuse with external distribution, ...
- Do I have permission from the copyrightholders to use the document? What is the scope of the permission? Territorial scope, duration, types of activities permitted: store, access by specific persons, access internally in the company, distribute externally, reuse without external distribution, reuse with external distribution, ...
- Can I claim blanket permission from a collective rights collecting society?

If copyright-protected documents are created within the company, the question arises who is the copyrightholder. A number of factors determine the answer to this question and these may differ in various jurisdictions. To illustrate, in some countries the copyright on works created by employees as part of their designated tasks automatically falls to the employer (e.g. in The Netherlands). In other countries, the copyright rests with the employee who actually created the work and transfer of copyright must be explicitely agreed upon (e.g. in Belgium). The corporate policy dealing with copyright should determine which factors need to be recorded for each relevant jurisdiction.

For copyright-protected documents transferred to the company, the main concern is usually knowing what use is permitted. In a business context, very few legal exceptions apply that allow certain kinds of use without obtaining permission. The general rule is that permission from the copyrightholder is required. The problem is that is not always easy to find out who the copyrightholder is. Initially, the copyrightholder is the author of the work. Very frequently, the creator transfers his rights in whole or in part to someone else or even to several persons at the same time. In turn these copyrightholders may transfer their rights to others. Since who is the rightholder is not static, it makes little sense to record it in the metadata of a document, without adding specific context information. The basis for the rightholder's claim to copyright ownership should be added, e.g. copyright law (creator), licence agreement or terms of employment. If available, information about the scope of the copyrightholders claim should be recorded, e.g. non-exclusive worldwide licence. Finally, it may be useful to add status information about the copyrightholders claim, especially if the claim is the object of a legal dispute. The time at which these elements were recorded should be included, as copyright ownership may change hands afterwards without being known to the organisation holding the work in its posession.

Tying documents to the copyright licences that govern their use should allow for a compliance analysis to be made whenever necessary. The corporate policy should determine if an analysis

must be made immediately upon receipt and if so which types of use must be addressed. The two main categories of use are reproduction and communication to the public.

It is very common for documents to contain contributions from various people. Either because documents are the result of a joint effort (ie co-authors of a study, team of engineers designing a ship, ...), or because existing material is repurposed for use in new works (ie modification of software, incorporating existing designs into the blueprints for a new ship, ...). Keeping track of who contributed what and under which (legal) conditions is an enormous challenge.

4.2.2 Preservation

Long-term preservation will generally require that a great number of documents are converted into new formats. Given the broad meaning most copyright laws give to the notion of reproduction, such conversion operations would in principle require permission from the copyrightholders (unless an exception applies). The question then is whether or not separate permission must be obtained for the conversion to be legitimate.

Additional complications arise when technical protection measures must be removed or circumvented to ensure long-term preservation and/or availability, or in case rights management information related to copyright is removed or manipulated in the process. Following the WIPO treaty on these matters, most countries have more or less outlawed circumvention of technical protection measures and removal of rights management information. It should be noted that such actions are now criminal offences in many jurisdictions.

4.2.3 Dissemination

The ultimate goal of preservation is to make documents available to authorized parties when required. Making copyrighted materials available electronically raises questions as to whether this an activity covered by copyright. The EU directive states that 'making available to the public of their works in such a way that members of the public may access them from a place and at a time individually chosen by them' is covered by copyright, just as traditional forms of communication to the public are. As a matter of principle, it doesn't matter how many people have access: one person, a few people, or the whole world. It is a matter of interpretation wether employees of a company are 'members of the public'.[87] Thus, it should be checked whether the permission to copy a document into the corporate archive also covers making it available electronically internally and/or externally.

[87] For instance in Belgium only use within the family is explicitly considered use 'not in public'.

This problem doesn't arise as such with copyrighted materials in tangible form, ie. paper books. Reading a book is not an act subject to the control of the copyrightholder. As a note, distribution of tangible copies – for sale, for rent or for loan – is generally governed by copyright.

4.2.4 (Re)Use

Mere 'consumption' (reading, viewing, hearing) of copyrighted materials was traditionally not regulated by copyright. Copyright was meant to ensure the author an exclusive right of exploitation of his work, not to control every form of use by the public of the work. As indicated above, the shift from paper to electronic documents thoroughly changed how copyright works. Displaying a digital object on screen necessitates at least one copy to be made in the computer's working memory. No matter how fleeting, this copy enough for copyright restrictions to apply. In theory, this would make viewing a website illegal unless the copyright holder granted a licence. Fortunately, broad exemptions were enacted to still allow for mere consumption. As always with exemptions, determining their limits may not be straightforward and differences between countries may exist.

Another matter entirely is reuse of copyrighted works to create other works, examples are translation, citation, compilation, ... It is impossible to list all the situations in which existing works are transformed in the process of creating new ones. It is equally impossible to give a general answer as to when such reuse is permissable. The goal of copyright metadata is not to anticipate future uses and determine if they are permitted. Once reuse is being considered, copyright metadata should provide the information necessary to determine efficiently under which conditions that reuse is permitted.

To give a simple example, if copyright metadata indicates that a work was created by an employee and that in the relevant jurisdiction copyright is automatically assigned to the employer, the question of reuse is quickly resolved. The same is true if the metadata indicates that copyright was signed over to the organisation. Without sufficient metadata an investigation has to be launched into who owns the copyright.

When copyrighted works are reused in the creation of other works, recording the relationship between the source and the resulting document is of great value. Depending on the circumstances, the conditions for use of the source influence the conditions for creation, preservation, dissemination and use of the resulting work.

4.2.5 Requirements

What follows is a list of requirements for a copyright compliance metadata schema. Each requirement is summarized in a short label, which will be used in the sections below discussing a number of existing copyright-related metadata schemas.

Be able to unambiguously identify documents What?

Be able to register what jurisdiction documents 'live' in. Where was the document created or where did I receive it? Determining relevant jurisdictions is very difficult in multinationals. Documents may be created/received in one country, be stored in another country and sent to any number of additional countries. Perhaps not absolutely ALL documents need to have metadata added about which jurisdictions it has been in, but the metadata model ought to offer the possibility. Jurisdiction(s)?

Be able to tag all documents that the company considers copyright protected. Already this is a question of interpretation and depends on the jurisdiction in which the question is posed. Copyrighted in jurisdiction?

Designs for a ship or an oil platform are obvious examples of copyright-protected materials. Reports and studies are in all likelihood copyright-protected as well. Much less clear is whether and under which circumstances ordinary business letters or e-mails would be copyright-protected.

Where collections of documents are tagged as being copyrighted and to be living in one or more jurisdictions, it should be possible to point to the relevant copyright regulation per jurisdiction.

For instance, the collection of documents tagged with EU country jurisdictions could point to the relevant EU directives on copyright.

This is only a first step in copyright compliance. Being able to point to copyright regulation – intermediated through corporate policies – only provides background information.

For individual documents, the relevant circumstances of their creation within the company or transfer to the company should be recorded, allowing a compliance analysis to take place whenever required.

Be able to record the origin of the document. For works created within the organisation, metadata should record who was involved in creation. Who created a copyrighted work is essential in determining the term of copyright protection since this depends on the date of death of the author(s). Origin: creator?

For documents created outside the organisation, investigating who was involved in its creation is often not cost-effective, unless in case of works made to order. Recording the source of the Origin: source?

document and the reason for its transfer, should be feasible in most cases and will provide a lead if later on it becomes necessary to fully investigate copyright ownership and use conditions.

Be able to record information about claims of copyright ownership laid to works

Who is the claimant? Does the initial creator still own the copyright or has it been transferred? Copyright Claim: who?

What is the basis for the claim? The law, agreements or verdicts may be the basis for the claim to copyright ownership. Copyright Claim: basis?

What is the scope of the claim? Claims can be limited in space and/or time, notably because the claimant only obtained a licence with a limited scope. Copyright Claim: scope?

Status of the claim? Is the claim under dispute? Copyright Claim: status?

Be able to record events pertaining to documents When a document was created, received, disseminated, reused is important for the application of copyright regulation and copyright agreements. Events?

Be able to link documents to copyright agreements about them. Copyright agreements may be found in individual licence agreements, in employee contractual clauser or also in blanket licences negotiated with collective rights collecting societies. Licence?

Being able to find agreements quickly, allows to determine more efficiently what my usage rights are: Do I have permission from the copyrightholders to use the document? What is the scope of the permission? Territorial scope, duration, types of activities permitted: store, access by specific persons, access internally in the company, distribute externally, reuse without external distribution, reuse with external distribution, ...

Be able to link to particular copyright exemptions In some case, copyright law itself grants permission to use copyrighted works under certain conditions. Being able to determine which exemption is invoked in which jurisdiction(s) allows to determine more efficent what my usage rights are: What is the scope of the exception: store, access by specific persons, access internally in the company, distribute externally, reuse without external distribution, reuse with external distribution, ... Exemption in jurisdiction?

Be able to link work processes to handling of copyrighted works Why?

Both copyright agreements and copyright exemptions often tie usage rights to particular work processes. For instance, under the EU Copyright Directive, certain archives may disseminate works but only for the purpose of research or private study by the recipient.[88]

Be able to record the relationschip between documents Where (portions of) works are reused Relation-ship?

[88] Art. 5 §3 n EU Copyright Directive.

to create new works, ideally this relationship would be recorded in metadata. If at any time questions arise whether the resulting work as such or a particular use made of it infringes copyright, such metadata provides valuable leads.

4.2.6 Existing metadata models for copyright

Many metadata models exist that deal in more or less detail with copyright issues. The focus is usually on particular transactions for the use of copyrighted content, thus describing only particular actions that are permitted. It comes as no surprise that copyright metadata has received attention from sectors as the recording industry, publishers and television and movie industries, as copyright underpins their current business model.

IFLA: Functional Requirements for Bibliographic Records The IFLA (International Federation of Library Associations and Institutions) published its 'Functional Requirements for Bibliographic Records' (FRBR) in 1997, describing an entity-relationship model of bibliographic records.[89] The aim was to help libraries design better cataloguing systems and enable more efficient exchange of metadata. FRBR itself is not designed to manage copyright information, but can serve as a fundamental building block.

In FRBR a distinction is made between the following entities:

WORK -> EXPRESSION -> MANIFESTATION -> ITEM

A work is a "distinct intellectual or artistic creation. There is no single material object one can point to as the work."[90] A work once existed as an idea in the mind of its creator(s) and in the model serves to group together everything that has sprung from that idea.

An expression is the "intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of such forms."[91]

A manifestation represents the "physical embodiment of an expression of a work. The entity defined as manifestation encompasses a wide range of materials, including manuscripts, books, periodicals, maps, posters, sound recordings, films, video recordings, CD-ROMs, multimedia

[89] IFLA STUDY GROUP ON THE FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS, *Functional Requirements for Bibliographic Records*, München, K.G. Saur, 1998

[90] IFLA STUDY GROUP ON THE FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS, *Functional Requirements for Bibliographic Records*, *op. cit.* (as in n. **??**), p. 16.

[91] IFLA STUDY GROUP ON THE FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS, *Functional Requirements for Bibliographic Records*, *op. cit.* (as in n. **??**), p. 18.

kits, etc. As an entity, manifestation represents all the physical objects that bear the same characteristics, in respect to both intellectual content and physical form."[92]

An item is a "single exemplar of a manifestation."[93]

An example may make these entities more clear.

An item is something you have in your archive, e.g. a particular Lord of The Rings Trilogy Extended Version DVD set in good condition. This item is an exemplar of a large series of identical items, all the Lord of The Rings Trilogy Extended Version DVD sets, which represent a manifestation. Apart from the LOTR Trilogy Extended Version DVD's, there is also the manifestion in the form of VHS video, and also manifestion which is the (unique) Production Master.

All these manifestations are representations of one expression, namely the Lord of the Rings Extended Version motion picture.

Related expressions are: LOTR cinema version and LOTR Sound track. All these expressions relate to one work, namely the LOTR as it was conceived of by Peter Jackson.

This work is related to the work of Tolkien, also known under the name Lord of the Rings. Tolkien expressed his conception of LOTR as a written text (expression). The initial manifestation was the unique author's manuscript. Since then, many manifestations have been published. Items of LOTR are available in most public libraries.

The FRBR model also defines the entities 'person' and 'corporate body'. Both of these entities can be tied to works, expressions, manifestations and items. This expresses what a person or corporation has done in relation to a work, e.g. create or realize. The metadata defined in the FRBR does not allow to record the scope of a person's rights or claims in relation to a work, manifestation, expression or item.

Finally, FRBR models the subject of works in 'concepts', 'objects', 'events' and 'places'. Of course, works may also take persons or corporate bodies, as well as other works, expressions, manifestations and even items as their subject.

The model describes what exists, but not why it exists. The perspective of the content producer is not addressed and there are no entities and relationships to express the responsabilities and rights in the creation of the expression/manifestation/items.

The strength of this model lies in the differentiation it makes between the levels of existence of records, from the highly abstract 'work' to the very concrete 'item'. These levels can be found

[92] IFLA STUDY GROUP ON THE FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS, *Functional Requirements for Bibliographic Records*, *op. cit.* (as in n. **??**), p. 20.

[93] IFLA STUDY GROUP ON THE FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS, *Functional Requirements for Bibliographic Records*, *op. cit.* (as in n. **??**), p. 20.

implicitely in copyright law's provisions. Using the FRBR model allows to attach copyright metadata to the right level and avoid unnecessary duplication. Also it allows for great precision in determining the relationship between different copyrighted records.

It should be noted however that the terms work and expression are used in the FRBR in a specific sense, which does not correspond completely to the meaning attached to these terms in copyright law. In copyright law a 'work' is not the concept in the author's head but it is an original idea which has been realized in a certain form. In other words, a 'work' is an expressed original idea. Copyright law gives the author the exclusive right to create identical copies of his work and to create adaptations of it. Using the terms of FRBR, copyright law attaches these rights to expressions, with consequences for the production of manifestations and items: the creator has the exclusive right to produce items of a manifestation (reproductions), as well as to produce variant manifestations (reproductions) and manifestations of adapted expressions of the initial manifestation (adaptations).

Besides its obvious uses as a foundation for copyright metadata schemas, de Oliveira Lima suggests FRBR could also serve as a model for describing legal norms and the relationships between them.[94]

Further reading

- IFLA Cataloguing Section: FRBR Review Group `http://www.ifla.org/VII/s13/wgfrbr/`
- TILLETT, BARBARA, *What is FRBR? A conceptual model for the bibliographic universe*, Washington D.C., U.S.A., Library of Congress, 2004 , `http://www.loc.gov/cds/downloads/FRBR.PDF`
- BEARMAN, DAVID et al., "A common model to support interoperable metadata", *D-Lib Magazine*, Vol. 5 1999, Nr. 1 , `http://www.dlib.org/dlib/january99/bearman/01bearman.html`
- IANNELLA, RENATO, "Digital Rights Management (DRM) Architectures", *D-Lib Magazine*, Vol. 7 2001, Nr. 6 , `http://www.dlib.org/dlib/june01/iannella/06iannella.html`
- DE OLIVEIRA LIMA, JOÃO ALBERTO, "An Adaptation of the FRBR Model to Legal Norms", In BIAGIOLI, CARLO, FRANCESCONI ENRICO SARTOR GIOVANNI (ed.), *Proceedings of*

[94] DE OLIVEIRA LIMA, JOÃO ALBERTO, "An Adaptation of the FRBR Model to Legal Norms", In BIAGIOLI, CARLO, FRANCESCONI ENRICO SARTOR GIOVANNI (ed.), *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art4.pdf`.

of the V Legislative XML Workshop, European Press Academic Publishing, 2007 , `http://www.e-p-a-p.com/dlib/9788883980466/art4.pdf`, *op. cit.* (as in n. **??**)

Indecs The Indecs project aimed to develop "a genre-neutral framework among rights-holders for electronic IPR trading so that companies which at present are record companies, film companies, book and music publishers can trade their creations in a coherent single marketplace." The project partners realised there was a gap to bridge between descriptive metadata (the expertise of cataloguing institutions) and rights metadata. In this light it is useful to see inhowfar it could serve as an extension of FRBR.

The principle of unique identification of entities is a cornerstone of interoperable metadata according to Indecs.[95] It should be possible to identify an entity whenever it needs to be distinguished (principle of functional granularity).[96] Indecs Metadata Dictionary contains references to a number of external identification systems.[97]

Indecs defines as its overarching concept 'creations', which are any product of human imagination and/or endeavour by one or more parties in which rights may exist. What?

Creations are broken down in a number of types:

INDECS term	Definition	FRBF term
Artefact	A creation which is a thing	/
Abstraction	A creation which is a concept	Work
Expression	An event which is a creation	Expression
Manifestation	An artefact containing an infixion of an expression	Manifestation
Item	A single instance of an artefact	Item
Format	An artefact on which an expression may be infixed to create a manifestation	/

> "The main function of these distinctions is that each of these different types of creation may give rise to a different intellectual property right; for example, in an audio CD there are separate rights in the physical product (manifestation), the recorded performances (expressions) and the songs performed (abstractions), and these each require distinct metadata at some point in the commerce chain. These rights have different values in different jurisdictions, and will commonly be owned or controlled by different people and organisations. While music is used as an example,

[95] RUST, GODFREY, BIDE MARK (ed.), *The indecs metadata framework, Principles, model and data dictionary*, Indecs, 2000 , `http://www.doi.org/topics/indecs/indecs_framework_2000.pdf`, p. 9.
[96] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 10.
[97] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 28.

> parallel situations exist for all other genres of creation. Without the clear structural distinctions of this kind, effective rights management is impossible."[98]

Through creation-to-creation roles, the model allows to express how existing creations were reused in new ones.[99] The act of reuse is an event in which specific agents participate, having inputs – existing creation(s) – and outputs – new creations.[100]

Relationship? Origin: creator?

INDECS goes a step further than describing what is. The model also captures certain actions in relation to them:

- people make creations
- people use creations
- people do transactions about creations.

Besides actions, the model also captures the fact that people may have certain claims over specific entities:

- people make intellectual property
- people use intellectual property
- people own rights in intellectual property

Making and using creations are modelled as events, creatingEvent and usingEvent respectively.[101] The place where the creation was made or used can be noted in the context element.[102]

Jurisdiction(s)?

Transactions, contained in agreements, are also a particular kind of event (see further).

The Indecs models provides a very refined set of elements to determine precisely what role an agent has in the making, using or doing transactions about creations.[103]

Origin: creator? Origin: source?

The strength of the Indecs model is that it doesn not attempt to 'summarize' all the intellectual property regimes of the world into it's own terms. In stead it designs a separate 'namespace' for these legal concepts. The namespace has two legal entities, ipTypes and ipRights. ipTypes is any category of creations that is covered by an intellectual property right anywhere in the world.[104]

[98] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 26.
[99] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 29.
[100] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 21 ff.
[101] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 21 ff.
[102] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 19.
[103] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 19.
[104] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 30.

ipType	Definition
Work	As defined by the Berne Convention for the Protection of Literary and Artistic work Works, the WIPO Copyright Treaty and the TRIPS Agreement
Performance	As defined by the International Convention for the Protection of Performers, performance Producers of Phonograms and Broadcasting Organisations (Rome Convention), the WIPO Performances and Phonograms Treaty and the TRIPS Agreement
CriticalOrScientificPublication	As defined by Art. 5 of the European Directive harmonising the term of protection of copyright and certain related rights Scientific Publication
GovernmentTextsBelgium	As defined by Art. 8 §2 Belgian Copyright Act

An ipRight is the "authority granted by law or international convention to do or to authorise another person to do a defined act to intellectual property"[105]

Copyrighted in jurisdiction?

In the model iprStatement forms the tie between people and the creations over which they hold intellectual property.[106] By way of IPR agreements, rights to intellectual property can be passed on to another person.[107]

Licence?

Agreements are concluded by parties, who were either invested with intellectual property rights by the law or by other agreements.[108]

Copyright Claim: who?

The terms of the agreement are modelled in the elements 'permission', 'requirement', 'prohibition', 'iprTransfer'.[109] This can be used to describe the scope of the agreement.

Copyright Claim: basis?

Copyright Claim: scope?

Because the aim of Indecs is to allow metadata from a variety of sources to interoperate, the reliability of the metadata is a concern. This is addressed by the defining 'assertions', being an event in which a party makes a claim of veracity about something.[110] Possibly the fact that a copyright claim is the subject of a legal dispute could be seen as a challenge of its veracity.

Copyright Claim: status?

The main goal of the Indecs model is to support automatisation of transactions regarding intellectual property rights in creations. "Rights transactions depend on a 'chain' of grants of rights and of permissions: this chain is established initially by law or statute, in what may be viewed

[105] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 15.
[106] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 32.
[107] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 32 ff.
[108] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 15 and 30 ff.
[109] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 33.
[110] RUST, *The indecs metadata framework, Principles, model and data dictionary*, *op. cit.* (as in n. **??**), p. 35 ff.

as the original binding agreement that confers rights to a person. Whether the laws are concerned with copyright, patent law or other forms of ip is unimportant for the operation of the framework."

The Indecs project is tailored to the needs for e-commerce in the creative industries. Of course the issue of who owns which rights in what creation is the same for a movie production company as it is in any other industry. Further research is required to figure out precisely how far the Indecs model can help towards copyright compliance and where extensions may be needed.

Is the entity ipTypes robust yet flexible enough to deal with all the intellectual property laws a multinational encounters and the changes they see over time? Where do corporate policies fit into the model, if at all, prescribing for instance when IPR agreements are permissible and what they should state? Is there a need for more clarification of the business context?

Additionally, further research should show inhowfar the Indecs model could be expanded for use in other legal domains, for instance compliance with data protection rules, enforcement of confidentiality agreements, etc.

Requirement	**Metadata element**	**Comment**
What?	Creation identifiers	
Jurisdiction(s)?	(event + role.context.place)	Event includes makingEvent and usingEvent.
Copyrighted in jurisdiction?	(ipType + role.context.place) + (ipRight + role.context.place)	A particular kind of creation (ipType) is recognized in a country as giving rise to copyright (ipRight) with a particular territorial scope.
Origin: creator?	Agent role: contributor, creator, modifier, excerpter, compiler, etc.	
Origin: source?	Agent role: disseminator	
Copyright Claim: who?	person, party	The law grants intellectual property to 'persons'. Agreements about the transfer of intellectual property are concluded by 'parties'.
Copyright Claim: basis?	IntellectualPropertyRight, agreement.output, iprStatement	
Copyright Claim: scope?	output of agreement (permission, requirement, prohibition, iprTransfer)	

Requirement	Metadata element	Comment
Copyright Claim: status?	assertion?	An assertion is an event in which a party makes a claim of veracity about something. The fact that a copyright claim is the subject of a legal dispute could be seen as a challenge of its veracity.
Events?	(event + role.context.time)	
Licence?	agreement	
Exemption in jurisdiction?	iprStatement?	According to the model, iprStatement describes the ownership of intellectual property right in a creation or the entitlement to agree its exploitation. One might consider encoding the reverse in iprStatements as well, namely the prohibition to disagree with certain use, which is essentially what a copyright exemption does.
Why?	/	
Relationship?	Creation-to-creation relation roles	

Further reading

- RUST, GODFREY, BIDE MARK (ed.), *The indecs metadata framework, Principles, model and data dictionary*, Indecs, 2000 , `http://www.doi.org/topics/indecs/indecs_framework_2000.pdf`, *op. cit.* (as in n. **??**) `http://www.doi.org/topics/indecs/indecs_framework_2000.pdf`.
- `http://xml.coverpages.org/indecs2rdd.html`.

Dublin Core Metadata Initiative The most well-known product of the Dublic Core Metadata Initiative is 'Simple Dublin Core', a set of 15 essential metadata elements to describe information resources with. Simple Dublic Core has been formally accepted by a number of standardization bodies:

- ISO Standard 15836-2003 of February 2003 [ISO15836]
- ANSI/NISO Standard Z39.85-2007 of May 2007 [NISOZ3985]
- IETF RFC 5013 of August 2007 [RFC5013]

> "The Dublin Core Metadata Element Set is a vocabulary of fifteen properties for use in resource description. The name 'Dublin' is due to its origin at a 1995 invitational workshop in Dublin, Ohio; 'core' because its elements are broad and generic, usable for describing a wide range of resources.
>
> The fifteen element 'Dublin Core' described in this standard is part of a larger set of metadata vocabularies and technical specifications maintained by the Dublin Core Metadata Initiative (DCMI). The full set of vocabularies, DCMI Metadata Terms [DCMI-TERMS], also includes sets of resource classes (including the DCMI Type Vocabulary [DCMI-TYPE]), vocabulary encoding schemes, and syntax encoding schemes. The terms in DCMI vocabularies are intended to be used in combination with terms from other, compatible vocabularies in the context of application profiles and on the basis of the DCMI Abstract Model [DCAM]."[111]

The elements are briefly reviewed in what follows. In the margin reference is made to the requirements listed above to identify which metadata elements may used to fulfill them in part or in whole.

Title: a name given to the resource

What?

- The title is an important element allowing to identify the information object. Various objects may share the same name, thus more information is often required to provide unique identification.

Identifier: an unambiguous reference to the resource within a given content (URL,ISBN...) What?

Creator: an entity primarily responsible for making the content of the resource (a person, an organization, or a service) Origin: creator?

- The creator is a potential copyright claimant, but not necesserily the current copyright holder. Duration of copyright is calculated as lifetime of the creator plus a fixed number of years. Knowing who the creator is, provides a clue on what the origin of the object is.

Publisher: an entity responsible for making the resource available (a person, an organization, or a service) Origin: source?

- The publisher is also a potential copyright claimant, but not necesserily current copyright holder. The term 'publisher' often has a quite narrow meaning in the law, it is unclear how

[111] `http://dublincore.org/documents/dces/`.

broad the term is meant here. Knowing who the publisher is, provides a clue on what the origin of the object is.

Contributor: an entity responsible for making contributions to the content of the resource (a person, an organization, or a service) Origin: creator?

- A contributor is a potential copyright claimant, not necesserily current copyright holder. Knowing who the contributor is, provides a clue on what the origin of the object is. Duration of copyright protection may alse depend on the lifetime of contributors, as is the case for creators.

Date: A date of an event in the lifecycle of the resource (form YYYY-MM-DD) Events?

- Some dates are relevant for the existence of legal claims, notable in the case of copyright.

- The following refinements exist for the element Date[112]
- **Date.created: Date of creation of the resource**
- **Date.dateCopyrighted: Date of a statement of copyright.**
  - = this has some relevance in the U.S. for instance
- **Date.issued: Date of formal issuance (e.g., publication) of the resource.**
  - = in Belgium, this date marks the start of copyright term for anonymous works

- To determine the term of copyright protection, a number of other dates should be recorded, which are not currently described by the Dublin Core Metadata Initiative:
  - *Date of death of the Creator*: date.deathCreator: copyright expires 70 years after death of the Creator
  - *Date of death of the Contributor*: date.deathContributor: copyright expires 70 years after death of the (last) Contributor

Rights: information about rights held in and over the resource Licence?

- Typically a Rights element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the rights element is absent, no assumptions can be made about the status of these and other rights with respect to the resource.[113]

[112] `http://dublincore.org/documents/usageguide/qualifiers.shtml`.
[113] `http://dublincore.org/documents/usageguide/elements.shtml`.

- The Rights element can be used to record whether there is a copyright notice on the resource.

- There are only a few refinements available for the Rights element: Access Rights and Licence
- **Rights.accessRights: Information about who can access the resource or an indication of its security status.**

- **Rights.licence: A legal document giving official permission to do something with the resource.**
- Using this refinement allows to point to a particular licence tied to an information object, e.g. Creative Commons, GPL or proprietary licence agreement.
- This element wasn't designed to document the entire legal history of an object, it is unclear how much it could describe.

- Can the Rights element be used to record legal claims anyone may state to have over the resource, e.g. claims deriving from trademark, defamation, tort, data protection, ...

Copyright Claim: who?

- Simple Dublin Core does not describe who owns – or at least claims – rights over a resource. This is solved in Qualified Dublin Core, which adds the element Rightsholder:
- **Rightsholder: A person or organization owning or managing rights over the resource.**
- This element was designed to capture copyright information, it is unclear how well it would serve to describe all other legal claims a person or organization may state on a resource.

Source: a reference to a resource from which the present resource is derived

- Derivative works of copyrighted materials require permission for their creation or legal exemption on copyright.

- There are no refinements defined for the Source element. As such, it doesn't appear possible to describe in Simple Dublin Core why the derivative was made (business process) or to document whether and where permission was obtained to do so.

Relation: a reference to a related resource

Relationship?

- **Relation.hasPart: The described resource includes the referenced resource either physically or logically.**
- **Relation.isPartOf: The described resource is a physical or logical part of the referenced resource.**

- **Relation.hasVersion: The described resource has a version, edition, or adaptation, namely, the referenced resource.**
  - A derivative has been made of this resource
- **Relation.isVersionOf: The described resource is a version, edition, or adaptation of the referenced resource. Changes in version imply substantive changes in content rather than differences in format.**
  - This resource is a derivative of .... Is there a way to record whether the derivative was made with permission or whether a compulsory licence applies?
- **Relation.hasFormat: The described resource pre-existed the referenced resource, which is essentially the same intellectual content presented in another format.**
  - A copy or derivative has been made of this resource. Is there a way to record whether the copy or derivative was made with permission or whether a compulsory licence applies?
- **Relation.isFormatOf: The described resource is the same intellectual content of the referenced resource, but presented in another format.**
  - This resource is a copy or derivative of .... Is there a way to record whether the copy or derivative was made with permission or whether a compulsory licence applies?
- **Relation.references: The described resource references, cites, or otherwise points to the referenced resource.**
  - Expresses a link between files. For instance to link a parody with an original, or a right of reply to incorrect or slanderous material. Is there a way to record whether the reference was made with permission or whether a compulsory licence applies?
- **Relation.isReferencedBy: The described resource is referenced, cited, or otherwise pointed to by the referenced resource.**
  - Expresses a link between files. For instance to link a parody with an original, or a right of reply to incorrect or slanderous material. For instance to link a parody with an original, or a right of reply to incorrect or slanderous material. Is there a way to record whether the reference was made with permission or whether a compulsory licence applies?
- **Relation.replaces: The described resource is supplanted, displaced, or superseded by the referenced resource. When establishing a chain of versions, where only one version is valid, the use of isReplacedBy and Replaces allows the relationship to be expressed and the user directed to the appropriate version. In this case, the reciprocal relationships are quite important.**
- **Relation.isReplacedBy: The described resource supplants, displaces, or supersedes the referenced resource.**

- **Relation.requires: The described resource requires the referenced resource to support its function, delivery, or coherence of content. In the case of IsRequiredBy and Requires, there is a clearer need to express the Requires relationship than the IsRequiredBy, though both can be useful. This relationship is most often seen in relationships between software and documents or applications and hardware and/or software requirements.**
- **Relation.isRequiredBy: The described resource is required by the referenced resource, either physically or logically.**

Subject: a topic of the content of the resource (keywords, key phrases, or classification codes)

Description: an account of the content of the resource (an abstract or a table of contents)

Type: the nature of genre of the content of the resource

Format: the physical or digital manifestation of the resource (media type or dimensions of the resource)

Language: a language of the intellectual content of the resource (RFC3066 and ISO639 recommended)

Coverage: the extent or scope of the content of the resource. Coverage will typically include spatial location (a place name or geographic co-ordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity).

- Tying a resource spatial and temporal data is the most common use made of this element. Using Coverage to specify jurisdiction is less common. It is unclear how the element Coverage (jurisdiction) should be interpreted. Does it state in which jurisdiction the resource was created or wich jurisdiction(s) claims applicability?
- It should be noted that Coverage describes the resource as a whole. Thus is would seem this element can not be used in conjunction with Rights de describe in which jurisdiction described Rights pertain.

Provenance: A statement of any changes in ownership and custody of the resource since its creation that are significant for its authenticity, integrity and interpretation. The statement may include a description of any changes successive custodians made to the resource. Provenance is a part of Qualified Dublin Core.

Dublin Core is designed to describe what an information object is. It is not designed to describe why it exists or how it was created. This explains the limitations on the metadata model's capability of recording all legal metadata needed for full compliance by the holder of an information resource.

Requirement	Metadata element	Comment
What?	Title, Identifier	
Jurisdiction(s)?	/	
Copyrighted in jurisdiction?	/	
Origin: creator?	Creator, contributor	
Origin: source?	Publisher, Provenance	The publisher is only one possible source, amongst many others. Inhowfar provenance is appropriate for all other sources is unclear.
Copyright Claim: who?	Rightsholder	Element part of Qualified Dublin Core
Copyright Claim: basis?		
Copyright Claim: scope?		
Copyright Claim: status?		
Events?	Date	A limited list of events.
Licence?	Rights.licence	It is unclear how much information about agreements can be included
Exemption in jurisdiction?	/	
Why?	/	
Relationship?	Source, Relation	

Further reading

- BEARMAN, DAVID et al., "A common model to support interoperable metadata", *D-Lib Magazine*, Vol. 5 1999, Nr. 1 , `http://www.dlib.org/dlib/january99/bearman/01bearman.html`, *op. cit.* (as in n. **??**)
- RUST, GODFREY, "Metadata: The Right Approach", *D-Lib Magazine*, 1998 , `http://www.dlib.org/dlib/july98/rust/07rust.html`
- DEKKERS, MAKX, WEIBEL STUART, "State of the Dublin Core Metadata Initiative", *D-Lib Magazine*, Vol. 9 2003, Nr. 4 , `http://dlib.org/dlib/april03/weibel/04weibel.html`

California Digital Library - CopyrightMD Schema CopyrightMD Schema `http://www.cdlib.org/inside/projects/rights/schema/` v. 0.9

> "From 2005 through 2006, RMG did an analysis of the functional requirements related to copyright metadata, identified key data elements for expressing copyright

> metadata, and formalized these elements in the copyrightMD XML schema. A beta version of the schema is now available, with documentation and usage guidelines."[114]

The overarching element of the CopyrightMD schema is Copyright. The attributes Copyright.status (e.g. copyrighted, public domain, unknown) and Publication.status (e.g. published, unpublished, unknown) must be filled in for the metadata record to be valid. All other elements are optional.

There are 7 groups of subelements available. Their use is explained in the User Guidelines.[115] In the margin reference is made to the requirements listed above to identify which metadata elements may used to fulfill them in part or in whole.

Creation: Wrapper element for information about the creation of the resource. Non-repeatable.

Events?

- Year.creation: Contains the year the resource was created. Non-Repeatable.

Jurisdiction(s)? Copyrighted in jurisdiction?

- Country.creation: Contains the name of the country in which the resource was created. Non-repeatable.

Creator

Origin: creator?

- Creator.corporate: Contains the name of a corporate entity responsible for creating the resource. Repeatable.
- Creator.person: Wrapper element for information regarding an individual responsible for creating the resource. Repeatable
  - Name
  - Year.birth:Contains the year of birth of an individual responsible for the creation of the resource. Non-Repeatable.
  - Year.death: Contains the year of death of an individual responsible for the creation of the resource.
- Note: General information. Repeatable.

Publication: Wrapper element for publication information regarding the resource. Non-repeatable.

Jurisdiction(s)?

[114] `http://www.cdlib.org/inside/projects/rights/record.html`. RMG is the Rights Management Group, see `http://www.cdlib.org/inside/groups/rmg/`.

[115] `http://www.cdlib.org/inside/projects/rights/schema/copyrightMD_user_guidelines.pdf`

- Country.publication: Contains the name of the country in which the resource was published. Non-repeatable.

Origin: source?

- Publisher: Contains the name of the publisher of the resource. Non-repeatable.

Events?

- Year.publication: Contains the year the resource was published. Non-repeatable

Events?

- Year.copyright: Contains the year the resource was copyrighted.

Events?

- Year.renewal: Contains the year the copyright for the resource was renewed. Non-Repeatable.
- Note: General information. Repeatable.

Rights.holder: Wrapper element for information about the rights holder for the resource and contact information for the rights holder or rights holder's designee.Non-repeatable.

Copyright Claim: who?

- Name: name of either an individual or a corporate entity identified as copyright holder for the resource. Repeatable.
- Contact: Use to provide relevant contact information, when available and not confidential, for the person or institution to whom a user should address questions regarding usage of or permissions regarding the resource. Repeatable
- Note: General information. Repeatable.

Notice: Contains the copyright notice as it appears on the resource. Non-repeatable.

Copyright Claim: basis?

Services: Wrapper element for information relating to services that might be offered relating to the resource, such as providing copies. Repeatable.

- Contact: Use to specify the service(s) available and contact information for the person or institution providing the service(s).
- Note: General information. Repeatable.

General.Note: Contains a general note regarding copyright information for the resource, only for information that cannot be accommodated in a more specific copyrightMD element.

The CopyrightMD schema was developed in the context of U.S. law, as a result some of the elements are useful only in the U.S. Some adaptations would be needed in order to make the schema usable in other jurisdictions. When analysing it from an international perspective, some obstacles very quickly become apparant.

To make things more clear, consider the following case:

The same work is in the collection of a U.S. and a Belgian organisation. The U.S. organisation uses CopyrightMD to describe copyright status of the work, while the Belgian organisation uses a derived schema called CopyrightMD_BE.

The problems start right in the beginning with the two attributes of the root element 'copyright'. The first attribute is called 'copyright.status' and it's possible values are[116] :

- copyrighted - Under copyright.
- pd - Public domain: No further information.
- pd_usfed - Public domain: US Federal document.
- pd_holder - Public domain: Item dedicated to the public domain by the rights holder.
- pd_expired - Public domain: Item in the public domain because of expiration of copyright based on U.S. law.
- unknown - Copyright status of the resource is unknown.

Some of these values are only applicable in the U.S. context, though this is only one symptom of a bigger issue. The assessment whether a work is under copyright is only valid in relation to a specific jurisdiction. Copyright terms are different in the U.S. and in Belgium, and are calculated according to very different rules.[117]

A similar problem arises with the attribute 'publication.status', which can have 3 possible values: 'published', 'unpublished' and 'unknown'. This may seem pretty straightforward, but the term 'published' has a specific legal meaning in the U.S., that does not correspond with it's meaning in other countries. In Belgium, for instance, publication is irrelevant in determining the scope of copyright or it's term. The Belgian organisation might be tempted to (mis)use this attribute to record whether a work was divulged or not, meaning that the author willingly released his work into the world and as such exercised his moral right of divulgation. Thus, the interpretation of 'publication.status' depends on the jurisdiction.

Users comparing the copyright status metadata delivered by the U.S. and Belgian organisation need qualifying information regarding jurisdiction to correctly interpret these attributes. The same is true in the case where these organisations would want to exchange copyright metadata. A simple solution might be to add a third attribute 'jurisdiction' with a country code as it's

[116] GROUP, RIGHTS DATA MANAGEMENT (ed.), *CopyrightMD User Guidelines, Version 0.9*, California Digital Library, 2006 , `http://www.cdlib.org/inside/projects/rights/schema/copyrightMD_user_guidelines.pdf`, p. 7.

[117] DEKEYSER, HANNELORE and LIPINSKI, TOMAS, "Digital Archiving and Copyright Law: A Comparative Analysis", *International Journal of Communication Law and Policy*, 12 2008 , `http://www.ijclp.net/12_2008/pdf/dekeyserlipinski.pdf`.

value. Perhaps these attributes should refer to a thesaurus defining possible values relevant for that jurisdiction. The adequacy of such a solution needs further investigation, for one because the relevant jurisdiction might be smaller or larger than a single country. Also, a lot of duplicate data would be found in metadata describing the copyright status of a work for more than one jurisdiction.

The publication element is of key importance in U.S. law for the calculation of the copyright term.[118] As indicated above, publication is of little consequence in Belgium.[119] This is no major problem, as the Belgian organisation might simply choose to omit this element or to include it simply as evidence of divulgation.

Some of the child-elements have very close ties to the U.S. context as well. The 'year.copyright' element contains the year the resource was copyrighted, typically based on a copyright notice on the resource itself.[120] This year has no relevance whatsoever in determining copyright term in Belgium. Likewise, the 'year.renewal' element is devoid of meaning in Belgium, as there never has been a renewal system in place. Both elements are optional and the Belgian organisation can omit them without trouble.

The creation element is also highly relevant in the U.S. when calculation the copyright term.[121] In Belgium, the date of creation ('year.creation') is not relevant, though the place of creation is ('country.creation'). Foreign works, created outside the EU, are protected just as Belgian works with the condition of reciprocity.[122] All of these elements are optional, allowing the Belgian organisation to only supply the place of creation.

The creator element plays a different role in U.S. than in Belgian copyright law. Under Belgian law, the creator of a work is always a physical person. A corporation can become rights holder but is never the creator. The Belgian organisation should ensure that the (optional) element 'creator.corporate' is never encoded for works created in Belgium. As is the case in the U.S. for personal creations, the year of death is crucial in determining the length of copyright protection.

When it comes to collaborative creations in particular, the role of a co-creator may be of great importance. For instance, Belgian law provides a list of contributors to audiovisual works who are presumed by law to be co-creators.[123]

[118] HIRTLE, PETER, *Copyright Term and the Public Domain in the United States 1 January 2007*, 2007 , `http://www.copyright.cornell.edu/training/Hirtle_Public_Domain.htm`

[119] With the exception of the neighboring right for the first publication of a never-before published public domain work. Art. 2 §6 BCA.

[120] GROUP, *CopyrightMD User Guidelines, Version 0.9*, *op. cit.* (as in n. **??**), p. 16.

[121] HIRTLE, *Copyright Term and the Public Domain in the United States 1 January 2007*, *op. cit.* (as in n. **??**)

[122] Art. 79 BCA.

[123] Art. 14 BCA.

Both the rights holder and services elements may contain very useful information, provided it is known at what time this information was valid. Information regarding an agent or a rights collecting society could be stored in one of these elements.

The copyright notice, contained in the notice element, served a particular purpose in U.S. law at one point. In Belgium, this information is useful, in the sense that the person or corporation mentioned in the copyright notice may be presumed to be the right holder by all third parties.

One piece of information that is currently not respresented in the schema is information regarding terms of acquisition. The organisation may obtain works in various ways, through contracts, as a consequence of legal mandates or perhaps a legal privilege. Often, the terms of acquisition have an impact on the way the works in question may be made accessible to users and used by them.[124] A Belgian archive might invoke the copyright exemption for cultural and scientific heritage when obtaining certain materials, this fact should be recorded seeing as granting access to the public is strictly limited by the law.

Even more complicated is the situation of a multinational organisation who would want to use CopyrightMD to record copyright information. CopyrightMD is currently not flexible enough to record metadata over different jurisdictions. For instance, as there may only be one publication block it is not possible to record that a resource was multiple times, specifically in different countries. Likewise, there can only be one copyright notice tied to a resource. Every resource is supposed to be created in one country, however more and more resources are the result of contributions from various countries. Selecting only one country as country of creation may be artificial and have little significance from a legal point of view.

CopyrightMD describes what a resource is, not why it exists or what the process of creation was. This limits the value of the schema for use in connection with living documents contained in a records management system.

Requirement	**Metadata element**	**Comment**
What?	/	CopyrightMD is not a standalone metadata schema, but used in addition to other schemas. Identification of the work is left to the other schema.
Jurisdiction(s)?	Country.creation, Country.publication	

[124] COYLE, KAREN, *Rights in the PREMIS Data Model*, Washington D.C., U.S.A., Library of Congress, 2006 , `http://www.loc.gov/standards/premis/Rights-in-the-PREMIS-Data-Model.pdf`, p. 25 ff.

Requirement	Metadata element	Comment
Copyrighted in jurisdiction?	Country.creation	
Origin: creator?	Creator	
Origin: source?	Publisher	
Copyright Claim: who?	Rights.holder	
Copyright Claim: basis?	Notice	A copyright notice may contain false or outdated information.
Copyright Claim: scope?	Notice	
Copyright Claim: status?	/	
Events?	Year.creation, Year.publication, Year.copyright, Year.renewal	
Licence?	/	
Exemption in jurisdiction?	/	
Why?	/	
Relationship?	/	

Further reading

- GROUP, RIGHTS DATA MANAGEMENT (ed.), *CopyrightMD User Guidelines, Version 0.9*, California Digital Library, 2006 , `http://www.cdlib.org/inside/projects/rights/schema/copyrightMD_user_guidelines.pdf`, *op. cit.* (as in n. **??**)
- COYLE, KAREN, "Descriptive metadata for copyright status", *First Monday*, 10 2005, Nr. 10 , `http://www.firstmonday.org/issues/issue10_10/coyle/index.html`

PREMIS

> "The Preservation Metadata: Implementation Strategies Working Group, convened by OCLC and RLG, initially developed the PREMIS data dictionary as a specification with the goal of creating an implementable set of 'core' preservation metadata elements, with broad applicability within the digital preservation community. "[125]

The Premis working group published its Data Dictionary in 2005, followed by a set of XML schemas to support their implementation.[126] Version 2.0 of the PREMIS Data Dictionary for

[125] `http://www.loc.gov/standards/premis/`

[126] COMMITTEE, PREMIS EDITORIAL, *PREMIS Data Dictionary for Preservation Metadata*, PREMIS, 2008 , `http://www.loc.gov/standards/premis/` p. 1.

Preservation Metadata was published on April 3d 2008. Notably the Rights Entity part of the standard was revised considerably.

The PREMIS data model defines five entities: Intellectual Entities, Objects, Events, Rights, and Agents. Each semantic unit defined in the Data Dictionary is a property of one of these entities.[127]

Intellectual Entity	A set of content that is considered a single intellectual unit for purposes of management and description: for example, a particular book, map, photograph, or database. An Intellectual Entity can include other Intellectual Entities; for example, a Web site can include a Web page; a Web page can include an image. An Intellectual Entity may have one or more digital representations.
Object (or Digital Object)	A discrete unit of information in digital form.
Event	An action that involves or impacts at least one Object or Agent associated with or known by the preservation repository.
Agent	Person, organization, or software program/system associated with Events in the life of an Object, or with Rights attached to an Object.
Rights	Assertions of one or more rights or permissions pertaining to an Object and/or Agent.

Instances of Objects, Events, Agents, and Rights statements are uniquely identified through 'Identifier' containers, which can refer to external identifier schemas.[128]

Identifiers are repeatable for Objects and Agents; they are not repeatable for Rights and Events.

Objects and Agents often have multiple identities in a global environment, and across systems, and therefore are likely to have multiple identifiers. Rights and Events are considered to have a context limited to a particular preservation repository, and therefore do not require multiple identifiers.[129]

[127] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 5.
[128] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 12.
[129] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 13.

What?

Premis does not distinguish Intellectual Entity into Works, Expressions and Manifestations as FRBR does. The PREMIS model considers the description of Intellectual Entities to be outside of its scope, choosing to focus only on their 'representations' in electronic form (Objects).[130]

Events?

The Events entity describes actions that occur in the course of preservation.[131]

Relationships?

Premis does not record relationships between Intellectual Entities, only between Objects. This is in line with the preservation focus of Premis and is useful for linking various representations or versions of the same Intellectual Entity to each other. The model does not exhaustively list all the possible relationships between Objects, only determines which metadata must be captured.[132]

Origin: creator? Origin: source? Copyright Claim: who?

The Premis model recognizes the importance of Agents in their relation to Events and Rights. However, only a means to identify the agent and a classification of agent type (person, organization, or software) is defined in the Data Dictionary. Where additional metadata is required, this must be provided by other schemas.[133]

Copyright Claim: basis? Licence?

The Rights entity can describe statements of rights and permissions. Rights are entitlements allowed to agents by copyright or other intellectual property law. Permissions are powers or privileges granted by agreement between a rightsholder and another party or parties.[134]

The revision of the Rights entity is summarized by Lavoie as follows:

> "Like its original version, the Rights entity in PREMIS 2.0 is intended to support an automated process that determines if a particular preservation-related action is permissible in regard to an Object or set of Objects within the repository, as well as to record important information about the permission. However, key differences exist between the old and new versions of the Rights entity. In PREMIS 2.0, the permissionStatement container is replaced by a new rightsStatement container, which can be used to express three forms of intellectual property rights: those established by copyright, those established by license, and those established by statute. The Rights entity defines metadata applicable to all three forms of rights statement, such as identifiers, the nature, scope, and characteristics of the rights granted to the repository, the Object(s) to which the rights apply, and the Agents responsible for granting or administrating the rights. In addition, the new Rights entity defines metadata specific to copyright-, license-, and statute-based intellectual property rights. The result

[130] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 8 and 22.
[131] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 10.
[132] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 10 and 13.
[133] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 11.
[134] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 157.

> is a deeper, more nuanced description of rights in a digital preservation context, yet one that preserves the earlier version's practical orientation toward automated processing."[135] .

CopyrightMD served as inspiration for the new rightsStatement container. However:

> "It should be noted that the proposed uses of copyrightMD and PREMIS rights are rather different. The copyrightMD schema is intended to document factual information to allow a human being to make an informed copyright assessment of a given work. The PREMIS rightsStatement is intended to allow a preservation repository to determine whether it has the right to perform a certain action in an automated fashion, with some documentation of the basis for the assertion."[136]

A rightsStatement containing copyright information records who owns intellectual property in the described object. The jurisdiction from which the copyright stems must be indicated. Legal disputes can be recorded in CopyrightNote.

Copyright Claim: who?
Copyright Claim: basis?
Copyrighted in jurisdiction?
Copyright Claim: status?

A rightsStatement containing licence information details which permission(s) the archive has to execute preservation activities (copying, migration, ...). The terms of the licence can be included. Whether or not the licence is under dispute could be recorded in LicenceNote

Licence?

A rightsStatement containing statute information details which permission(s) a law or statute grants to the archive to execute preservation activities. The jurisdiction from which the statute stems must be indicated. Legal disputes can be recorded in StatuteNote.

Copyright Claim: scope?
Copyright Claim: status?
Exemption in jurisdiction?
Copyrighted in jurisdiction?
Copyright Claim: status?

As many rightsStatements as necessary can be tied to any object in the archive, thus making it possible to take into account the perspective of more than one legal system. The Rights container is extensible, allowing organisations to record additional metadata as needed.[137]

Unlike Indecs, PREMIS 2.0 does not model the transactions about objects, only the resulting permissions. However, PREMIS 2.0 provides hooks for extensions to the rights entity to be developed.

[135] LAVOIE, BRIAN F., "PREMIS With a Fresh Coat of Paint. Highlights from the Revision of the PREMIS Data Dictionary for Preservation Metadata", *D-Lib Magazine*, Vol. 14 2008, Nr. 5/6 , `http://www.dlib.org/dlib/may08/lavoie/05lavoie.html`

[136] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 11.

[137] COMMITTEE, *PREMIS Data Dictionary for Preservation Metadata*, *op. cit.* (as in n. **??**) p. 19.

Requirement	Metadata element	Comment
What?	Identifier	
Jurisdiction(s)?	/	Could be implemented through Rights-Extension.
Copyrighted in jurisdiction?	Rights.copyrightJurisdiction, Rights.statuteJurisdiction	When a rights container is recorded with 'copyright' or 'statute' as its basis, then the jurisdiction from which the right stems is recorded.
Origin: creator?	Agent, Rights.linking-AgentRole	
Origin: source?	Agent, Rights.linking-AgentRole	
Copyright Claim: who?	Agent, Rights	
Copyright Claim: basis?	Rights	Copyright, Licence or Statute.
Copyright Claim: scope?	Rights.licenceTerms, Rights.rightsGranted	
Copyright Claim: status?	Rights.copyrightNote, Rights.licenseNote, Rights.statuteNote	
Events?	Event	
Licence?	Rights.license	
Exemption in jurisdiction?	Rights.statute	
Why?	/	
Relationship?	Relationship	

Further reading

- COMMITTEE, PREMIS EDITORIAL, *PREMIS Data Dictionary for Preservation Metadata*, PREMIS, 2008 , `http://www.loc.gov/standards/premis/`, *op. cit.* (as in n. **??**)
- LAVOIE, BRIAN F., "PREMIS With a Fresh Coat of Paint. Highlights from the Revision of the PREMIS Data Dictionary for Preservation Metadata", *D-Lib Magazine*, Vol. 14 2008, Nr. 5/6 , `http://www.dlib.org/dlib/may08/lavoie/05lavoie.html`, *op. cit.* (as in n. **??**)
- COYLE, KAREN, *Rights in the PREMIS Data Model*, Washington D.C., U.S.A., Library of Congress, 2006 , `http://www.loc.gov/standards/premis/Rights-in-the-PREMIS-Data-Mo` `pdf`, *op. cit.* (as in n. **??**)

Rights Expression Languages As part of the development of technical protection measures[138] to control the use of content by the public, a number of projects to develop 'rights expression languages'

Most of these Rights Expression Languages assume that the basis for use is a specific agreement between the content provider and the user. What is expressed is who can do what with which files (play, store, excerpt, ...).

From the perspective of enforcing compliance in an company, Rights Expression Languages may provide a useful building block. Just like media companies aim to exert control over users in an effort to protect their intellectual property, companies may wish to exert control over the use of records to enforce compliance with a range of legal requirements, such as intellectual property, confidentiality, data protection.

The aim of Rights Expression Languages is to tell a device whether or not to allow the use of certain content by a user. Determining why the user should or should not be allowed the requested form of use is outside the scope of REL.

Going into detail into all the REL under development is beyond the means of this study. What follows is a list of prominent examples.

- Open Digital Rights Language `http://www.odrl.net/` (ODRL)
- XrML `http://www.xrml.org/`
- MPEG-21 Part 5 Rights Expression Language (REL)

Further reading

- `http://xml.coverpages.org/drm.html`

BBC SMEF An industry that has a long history of dealing with production, use and reuse of copyrighted material is the audiovisual industry. Perhaps surprisingly, there is no generally accepted practice on how to go about rights management, nor is there a standard for copyright metadata.

Obviously, encoding the copyright notice featured on a movie, documentary or other audiovisual work into a metadata field is a trivial exercise. Capturing who all the contributors were in the

[138] This is more commonly known as DRM or Digital Rights Management, a term which will not be used here to avoid confusing with Digital Records Management. Digital Rights Management is a misleading term to apply to most common technical protection measures, as they do not deal with legal rights (who owns intellectual property rights or contractual to what) but with use restrictions.

creation of an audiovisual work, from the script writer, to the author of song track lyrics or the designer of artwork in the background, is far from trivial. Going one step further to also capture the legal grounds for inclusion of that contribution into the final work as well as its exploitation is a daunting task.

The BBC has taken on this challenge in the development of its 'Standard Media Exchange Framework', which is aimed to "support and enable media asset management ('MAM') as an end-to-end process across its business areas, from commissioning to delivery to the home."[139]

The data model is able to accompany a production from it's first conception as an idea over the production phase until it's broadcast and transfer. There are over 270 entities defined in the model.

A cluster of entities describes television programmes at varying levels of granularity (objects). A programme can be divided into items and shots, and it can be grouped with other programmes.

The entity 'REPORTING_COPYRIGHT' gives details on contributions and source material used in programs (such as scripts, music, or even extracts from other programs).

The entity 'role' provides the answer to the question "Who does what under which conditions?" by linking together parties (which can be either persons or organisations) with objects, and with contracts (or contract line). Thus, a role could be that of a script author (linking a person, her contract for work, and the respective script), or that of an actress (linking a person, her employment contract, and the respective program).

To give an example, if an Andy Warhol painting is used in a programme item an instance of 'REPORTING_COPYRIGHT' is created to describe the Andy Warhol painting and linked with the item. An instance of 'ROLE' is created linking the object described in 'REPORTING_COPYRIGHT' (the painting) with an organisation (the Warhol Estate) and a contract (terms of permission for reuse).

The entity 'contract' contains one or many 'clauses', or inother words all the statements used in a contract. The clauses should detail all the contractual terms, for instance the broadcaster's rights with respect to the material.

The model does not have room for obligations or rights stemming from sources other than agreements, notably regulations. Thus reuse of copyrighted material permitted by an exemption in copyright law could not be registered in the SMEF data model.

One might think that the entity 'right' would be suited for this purpose, especially in light of its broad definition as "[a]n interest, or permission, which is recognised and protected by law".[140]

[139] `http://www.bbc.co.uk/guidelines/smef.shtml`
[140] SMEF Data Model, p. 284.

However, upon close inspection, the right meant here only covers transmission rights, and not other types of legal interests such as copyright or privacy. This observation is based on a systematic interpretation, noting that the description of the entity 'RIGHT_TYPE' states that examples of types of right are the right to broadcast or the right to publish (both being categories of transmission)[141]; and taking into consideration the attribute 'RIG_Publication_Count', which indicates the total number of transmissions acquired by an outlet.[142]

The SMEF Data Model is an example of how to make the connection between business process (audiovisual production and broadcast) and legal information. Not only are contracts included in the model, but legally relevant activities – reuse of copyrighted material – is annotated.

[141] SMEF Data Model, p. 287.
[142] SMEF Data Model, p. 284.

Chapter 5

Recordkeeping

Organisations and their records managers have a number of sources to turn to in developing their recordkeeping approach. Legal requirements inevitably impact creation, maintenance and preservation of records. Other than that, inspiration can be found in standards, guidelines and best practices.

5.1 Standards

Since 1997, a subcommittee exists within ISO which has records and archives management as its domain, Subcommittee (SC) 11 records and archives management, part of ISO Technical Committee (TC) 46 Information and documentation.

This subcomittee has published a number of standards, amongst which ISO 15489 Information and Documentation – Records Management, the first international records management standard. Another is ISO 23081 Metadata for records.[1]

ISO 15489 is divided in two parts:

- Part 1: General (ISO 15489-1:2001)
- Part 2: Guidelines [Technical Report] (ISO/TR 15489-2:2001)

ISO 15489-1 defines metadata as "data describing context, content, and structure of records and their management through time."

[1] For a description of the work of the subcommittee and these standards see HOFMAN, HANS, "Developments in ISO standards for recordkeeping", 2005 , `http://dlmforum.typepad.com/Paper_HansHofman_onstandards.pdf`.

ISO 15489 Part 2 statest that "a records management policy statement is a statement of intentions. It sets out what the organization intends to do and, sometimes includes an outline of the program and procedures that will achieve those intentions."

A records management policy should include:[2]

- Purpose, scope and applicability of the policy
- Rules and responsabilities
- Ownership, legal status, access rights and privacy
- Goals, principles and objectives
- References to documentation and related policies.

A number of European initiatives regarding functional requirements for records management systems are reviewed in WALDRON, MARTIN, "Adopting electronic records management: European strategic initiatives", *The Information Management Journal*, 2004, Nr. July/Aug. Amongst others Moreq, Domea, Afnor standard NZ 42-013 are explained briefly.

In Canada, the 'Electronic records as documentary evidence' (CGSB 72.34) standard was issued on December 1st 2005. The standard establishes requirements for the creation of electronic records in any form to ensure that their authenticity can be demonstrated. It is important to note that, despite the name, compliance with the standard will not by itself guarantee that electronic records will be accepted by courts as evidence. This decision is the sole competence of the courts.[3] The standard does maximize the probability of admissibility of electronic records.[4] The standard is tailored to the legal rules of evidence in Canada, but many requirements will no doubt be just as relevant in other jurisdictions.

5.2 Best practices and guidelines

A number of organisations have issued guidelines on the subject. Notable examples of the latter are the following:

- Guide for Managing Electronic Records from an Archival Perspective `http://www.ica.org/en/node/30019`

[2] MYLER, ELLIE, "Minimizing Risks through a corporate information compliance initiative", *The Information Management Journal*, 2008, Nr. Jan/Feb, p. 59.
[3] **?**, , p. 39.
[4] **?**, , p. 40.

- DIRKS `http://www.records.nsw.gov.au/recordkeeping/dirks-manual_4226.asp` (Strategies for Documenting Government Business)
- VERS `http://www.prov.vic.gov.au/vers/vers/default.htm` (Victorian Electronic Records Strategy)

There are so many guidelines out there that no attempt is made here to provide a comprehensive list. All of these guidelines are more or less abstract in nature, therefor it would be of great interest to know how they are implemented in practice by various organisations. Unfortunately, precious little information is available on just how organisations manage their records.

In Switzerland, the Association of Swiss Archivists[5] conducted a Records Management Survey Schweiz in selected sectors of the private sector in 2006.[6] The survey concluded that in practice records management is often limited to traditional written records in particular financial records and personel files. Digital born documents often escape centralized records management based on selection and maintenance plans. There is a great need for training of all employees to improve records management in companies.

The Virginia Commonwealth University (VCU) conducted a study on how data management is practiced worldwide: "Measuring Data Management Practice Maturity: A Community's Self-Assessment".[7] The study concludes that most organizations do not manage information well. The study focusses on the quality of data, mostly held in databases, not on records management as such. Thus the relevance of this work for records management is not entirely clear.

Developing a records management requires involvement from the various stakeholders in an organisation.

> "While managing paper records has been the undisputable domain of records managers, managing electronic records requires teamwork. Today, it is common to see representatives of information technology, business units, records management, and legal services or compliance coming together to develop policies and procedures to address electronic records and information management issues. As part of the team, records managers are in a position to bring tbeir expertise in categorizing, classifying, and indexing documents to the development of metadata."[8]

The questions to be answered are:

[5] `http://www.vsa-aas.org/`.
[6] `http://www.vsa-aas.org/de/aktivitaet/earchiv/taetigkeit/rmsurvey/`.
[7] AIKEN et al., *Computer 40 [2007]*.
[8] FRANKS and KUNDE, *The Information Management Journal Sept/Oct [2006]* (as in n. **??**), p. 56.

- What records do we have? (Records manager, business unit, legal)
- Which records should we have? (Legal, business unit, Records manager)
- Why do we make/preserve these records? (Legal, business unit)
- How do we make these records? (business unit, IT)
- How should we make these records? (Legal)
- How do we preserve these records? (Records manager, IT)
- What is the budget? (Management)

An important part of the recordkeeping policy is the records retention schedule. A strategic overview of how to develop a retention schedule in a large organisation is described by Fischer.[9] Important advice is to thoroughly document the process, such as to enable records managers to determine why specific retention periods were decided for the various types of records. In case of a dispute, or generally when retention periods are questioned, such documentation provides valuable information.[10]

Establishing a coherent retention schedule is a particularly difficult exercise for multi-national companies. A small example is given in JONES, THOMAS M. et al., "Going global. Mapping an international records retention strategy." *The Information Management Journal*, 2008, Nr. May/June. After an overview of the most important U.S. retention requirements, the following best practice is described:

> "Absent other guidance, follow U.S. rules – In cases where a particular type of record is not governed by foreign retention requirements, multinational organizations should simply adhere to their current U.S. retention periods as their global default standard.
>
> Adopt global norms where they exceed U.S. practice – Given that some international records retention guidelines are longer than U.S. guidelines, the best practice is to adopt the longer retention period. For example, the prevailing U.S. practice for retaining certain tax records is seven years, whereas the minimum retention period in Germany is 10 years. For excessively long single-country requirements, issue "exception" policies –
>
> For excessively long retention requirements of a single country, make reasoned decisions as to compliance. For example, Argentina and Puerto Rico require accounting records to be retained until closure of business plus 10 and five years,

[9] FISCHER, LAURIE, "Condition critical: developing records retention schedules", *The Information Management Journal*, Jan/Feb 2006. See also MYLER, *The Information Management Journal 2008* (as in n. **??**).

[10] FISCHER, *The Information Management Journal Jan/Feb [2006]*, *op. cit.* (as in n. **??**), p. 34.

> respectively. In these cases, the best strategy is to issue singie-country "exception" retention policies that mandate compliance for the business operations located there."[11]

The article by Jones e.a. conveniently omits the problems caused by differing legal systems imposing radically conflicting obligations. Nowhere is this more apparant than with respect to privacy questions. To give but one example, in the U.S. organisations are relatively free to monitor their employees e-mail communications[12] , whereas their European counterparts must take into account privacy laws. Data protection legislation is a recurring source of conflict for companies operating in the U.S. and in Europe, as was demonstrated recently in the SWIFT case.

Apart from the question how long a record should be preserved, are questions regarding the level of protection – or conversely access – to be given to records, who is the custodian responsible for the record,[13] and how quickly a record must be produced.

5.3 Regulatory recordkeeping requirements

5.3.1 Public Sector

CH

The Swiss ISB (Informatikstrategieorgan Bund) which coordinates the implementation of e-Government on the federal, cantonal and municipal level is developing the GEVER project, which stands for Geschäftsverwaltung (business administration).[14] The aim is automating work processes troughout government agencies on all levels. Quite a number of regulations impact the creation, management and long term preservation of records by the government, as a study of the Swiss Federal Archives revealed.[15] Another document lists the standards relevant to the GEVER project.[16]

[11] JONES et al., *The Information Management Journal 2008*, *op. cit.* (as in n. **??**), p. 36.

[12] This liberty is taken for granted, see for instance PEGLAR, ROB, "Evidence management solutions for mitigating e-records risks", *The Information Management Journal*, 2007, Nr. July/Aug.

[13] FISCHER, *The Information Management Journal Jan/Feb [2006]*, *op. cit.* (as in n. **??**), p. 26.

[14] `http://www.isb.admin.ch/themen/architektur/00078/index.html?lang=de`.

[15] BUNDESARCHIV, SCHWEIZERISCHES (ed.), *Rechtliche Aspekte elektronischer Geschäftsverwaltung*, Bern, 2007 , `http://www.isb.admin.ch/themen/architektur/00078/00080/00212/index.html?lang=de\download=NHzLpZeg7t,.`

[16] EFD, EIDGENÖSSISCHES FINANZDEPARTEMENT (ed.), *Vorgaben zu GEVER / Records Management*, 2007 , `http://www.isb.admin.ch/themen/architektur/00078/00197/00214/index.html?lang=de\download=NHzLpZeg7t,.`

SI

The Slovenian Protection of Documents and Archives and Archival Institutions Act (PDAAIA)[17] contains provisions regarding the preservation of documents both in physical and electronic form. The law applies to the public and private sector.[18]

U.S.

FDA part 11 Electronic Records and signatures

In March of 1997, the FDA issued regulations determining under which circumstances electronic records and electronic signatures are accepted as trustworthy, reliable and equivalent to paper records and handwritten signatures executed on paper.[19] The aim of these regulations was to make allowance for widespread use of electronic technology, in a way compatible with the FDA's responsibility to protect the public health. Subsequently, the FDA published a compliance policy guide[20], issued numerous (draft) guidance documents covering such topics as validation of computer systems time stamps, maintenance and copying of electronic records, as well as a glossary of terms.[21] The FDA opted for a phased approach, progressively increasing the number of record types that may be submitted in electronic form.[22] Additionally, entities seeking to submit records electronically must first consult with the intended receiving unit of the agency on all practical issues concerning electronic submission, ie method of transmission, media, file format, technical protocols.[23] Records required to be maintained but not submitted could be kept in electronic form immediately from entry into force of the rules, provided that the requirements are met.[24]

Through extensive contact with the industry subject to FDA regulations, a number of concerns were raised regarding the regulation on electronic records and electronic signatures, notably

[17] See `http://www.arhiv.gov.si/en/archival_regulations_and_standards/` for an English translation of the act.

[18] See HAJTNIK, *DLM Forum Meeting, Ljubljana, 8-9 april 2008 2008* (as in n. **??**).

[19] These rules were incorporated in the Code of Federal Regulations, Title 21 Food and drugs, Chapter I Food and drug administrations, department of health and human services, part 11 Electronic records; electronic signatures, Federal register, Vol. 62, Nr. 54, March 20, 1997, p. 13430 ff., (hereafter 21 CFR Part 11). See Section 11.1 of 21 CFR Part 11 for an outline of the scope of these rules.

[20] Sec. 160.850: Enforcement Policy: 21 CFR Part 11 (CPG 7153.17), introduced 13/05/1999, revoked on 19/02/2003 (68 Fed. Reg. 8775).

[21] Federal Register, Vol. 68, nr. 37, February 25, 2003, p. 8776.

[22] Section 11.2(b)(1) of 21 CFR Part 11.

[23] Section 11.2(b)(1) of 21 CFR Part 11.

[24] Section 11.2(a) of 21 CFR Part 11.

with regard to cost of compliance, resulting restrictions on use of technology and discouragement of technical innovation. This has led the FDA to announce a complete re-examination of 21 CFR Part 11.[25] This announcement was accompanied by the decision to withdraw the existing draft guides and the compliance policy guide. Though 21 CFR Part 11 remains in force, the FDA has decided to exercise discretion with regards to enforcement, as detailed in the final guidance document[26] posted on their website on March 9th 2003. In particular discretion will be exercised regarding to validation, audit trail, record retention, and record copying requirements. All other provisions of 21 CFR Part 11, in particular those concerning controls for open and closed systems and requirements for electronic signatures, will be enforced as before.[27]

21 CFR Part 11 rules require that procedures for electronic records management are established and followed to ensure the authenticity, integrity and, where appropriate, the confidentiality of electronic records is guaranteed. Also, the signer must not be able to easily repudiate any signed record as not being genuine.[28] The rules break down this general aim into a number of requirements, both of an organizational and technical nature. Each organisation is responsible for drawing up their own procedures and designing their systems to meet these requirements, as well as for keeping them up to date.

Procedures should describe[29]:

- How electronic records will be maintained;
  - Validation of systems to ensure accuracy, reliability, performance and integrity.[30]
  - Accessibility of accurate and complete copies in human readable and electronic form throughout the records retention period [31]
  - Use of authority checks[32]
- Storage conditions and precautions;
  - Use of secure, time-stamped audit trails documenting the life-cycle of each record.[33]
  - Use of operational system checks to enforce permitted sequencing of steps and events.[34]

[25] Federal Register, Vol. 68, nr. 37, February 25, 2003, p. 8775.
[26] FDA, "Part 11, Electronic Records; Electronic Signatures – Scope and Application", available at `http://www.fda.gov/cder/guidance/5667fnl.htm`; as announced in Federal Register, Vol. 68, nr. 37, February 25, 2003, p. 8776.
[27] Federal Register, Vol. 68, nr. 37, February 25, 2003, p. 8776.
[28] Section 11.10 of 21 CFR Part 11.
[29] Draft Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures Maintenance of Electronic Records, v. July 2002 (withdrawn), p. 7.
[30] Section 11.10(a) of 21 CFR Part 11.
[31] Section 11.10(b) and (c) of 21 CFR Part 11.
[32] Section 11.10(g) of 21 CFR Part 11.
[33] Section 11.10(e) of 21 CFR Part 11.
[34] Section 11.10(f) of 21 CFR Part 11.

  - Use of device checks to ensure source validation.[35]
  - Maintenance of systems documentation[36]
- Retrieval and access restrictions;
  - Limiting system access to authorized individuals[37]
- The technical approach to long term electronic record storage; and,
- Personnel responsibilities for relevant tasks.
  - Education and training of personnel[38]
  - Accountability and responsability of individuals for any actions initiated under their electronic signature in accordance with written policies[39]

When open systems are used – meaning an environment in which system access is not controlled by persons who are responsible for the content of electronic records that are on the system – additional measures to ensure record authenticity, integrity and confidentiality may be necessary. The rules cite document encryption and digital signature techniques as examples.[40]

Factors that could potentially affect the reliability of electronic records during their records retention periods should be identified and controlled, otherwise information that the electronic records should convey might not be complete, accurate, or usable[41]:

- Data encoding
- Metadata
- Media
- Hardware
- Software
- Viewers

Only electronic signature methods that provide strong authentication are acceptable for use in matters controlled by the FDA. Organizations must verify the individual's identity before assigning that person an electronic signature or the means to create electronic signatures.[42] The electronic signature method employed shall ensure that only the genuine owner can use it to

[35] Section 11.10(h) of 21 CFR Part 11.
[36] Section 11.10(k) of 21 CFR Part 11.
[37] Section 11.10(d) of 21 CFR Part 11.
[38] Section 11.10(i) of 21 CFR Part 11.
[39] Section 11.10(j) of 21 CFR Part 11.
[40] Section 11.30 of 21 CFR Part 11.
[41] Draft Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures Maintenance of Electronic Records, v. July 2002 (withdrawn), p. 8.
[42] Section 11.100(b) of 21 CFR Part 11.

generate signatures.[43] Likewise, electronic signatures must be unique to an individual.[44] For the purpose of legal certainty, organizations using electronic signatures must certify – on paper – to the FDA that they intend these signatures to be legally binding.[45] Additional requirements apply to electronic signatures not based upon biometrics:[46]

- Signatures must be based on two distinct identification components, such as an identification code and a password;
- Attempted use of such signatures by anyone other than the owner, shall require collaboration of at least two people;

When identification code and password systems are in place, the organization must implement controls to ensure their security and integrity[47]:

- the combination of code and password must be unique;
- the security of the combinations must be checked periodically;
- loss management procedures must be in place;
- prevention measures against unauthorizes use must be in place;
- testing procedures for devices that bear or generate identification code or password information;

The regulation specifies a number of requirements for 'signature manifestations'. The printed name of the signer, the date and time of signature and the meaning of the signature (review, approval, responsibility or authorhip) shall be associated with the signature and made visible as part of any human readable form of the electronic record, either on a computer display or on a printout.[48] Also, electronic signatures must be linked to electronic records such that they cannot readily be excised, copied or transferred to create false electronic records.[49]

The draft guidance on maintenance of electronic records detailed an approach for migration of digitally signed records.[50]

A major element in the procedures to ensure authenticity and integrity of electronic records are the production of secure, computer-generated, time-stamped audit trails which independently

[43] Section 11.200(a)(2) and (b) of 21 CFR Part 11.
[44] Section 11.100(a) of 21 CFR Part 11.
[45] Section 11.100(c) of 21 CFR Part 11.
[46] Section 11.200(a)(1) and (3) of 21 CFR Part 11.
[47] Section 11.300 of 21 CFR Part 11.
[48] Section 11.50 of 21 CFR Part 11.
[49] Section 11.50 of 21 CFR Part 11. See also the explanatory notes preceeding the rules, Federal register, Vol. 62, Nr. 54, March 20, 1997, p. 13455 ff.
[50] Draft Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures Maintenance of Electronic Records, v. July 2002 (withdrawn), p. 20 ff.

record the date and time of operator entries and actions that create, modify, or delete electronic records.[51] The accuracy of time-stamps was addressed further in the draft guidance on time-stamping[52]

Not only should procedures be developped to address all the above points, the computer systems used in the implementation of an electronic records and electronic signature system must be validated to ascertain whether the goals are met consistently. A draft guidance was issued on validation of computer systems, but later withdrawn.[53] The draft guidance listed a number of key principles concering systems validation. Establishing and documenting system requirements specifications is a necessary first step in validation as this serves as the baseline for benchmarking. Through validation evidence must be obtained that the

computer system implements the requirements consistently. Thorough documentation of the validation process is of great importance in order for it to serve its purpose. Validation documentation should include a validation plan, validation procedures, and a validation report, and should identify who in management is responsible for approval of the plan, the procedures and the report.[54] The FDA observes that objective self-evaluation is difficult, thus , where possible, and especially for higher risk applications, computer system validation should be performed by persons other than those responsible for building the system.[55] Once a system is validated, change control is necessary to monitor if and when new validation actions are required.[56]

The Draft guidance acknowledges that the internet may play a role in electronic recordkeeping, notably for transfer of records. While validating the internet per se is not possible, the systems at either end ought to be validated.[57]

The FDA rules on electronic records and electronic signatures provide a useful overview of requirements to be met for the preservation of authentic and reliable records. Implementation is left up to the organizations themselves.

Further reading

[51] Section 11.10(e) of 21 CFR Part 11.

[52] Draft Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures; Time Stamps. v. February 2002.

[53] Federal Register, Vol. 68, nr. 37, February 25, 2003, p. 8776.

[54] Draft Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures; Validation, v. August 2001, p. 6.

[55] Draft Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures; Validation, v. August 2001, p. 10.

[56] Draft Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures; Validation, v. August 2001, p. 10.

[57] Draft Guidance for Industry 21 CFR Part 11; Electronic Records; Electronic Signatures; Validation, v. August 2001, p. 13 ff.

- FDA part 11 `http://www.fda.gov/ora/compliance_ref/part11/`
- `http://www.21cfrpart11.com/index.html`

Government Paperwork Elimination Act

Under the Government Paperwork Elimination Act (GPEA), Pub. L. No. 105-277, 1701-1710 (1998) (codified as 44 U.S.C.A. 3504 n. (West Supp. 1999)), Federal Executive agencies are required to implement infrastructure to allow for electronic maintenance, submission, or disclosure of information as a substitute for paper whenever practicable, as well as for the use and acceptance of electronic signatures. Guidance to assist agencies in implementing GPEA's requirements was developed by the Office of Management and Budget (OMB) in "Procedures and Guidance; Implementation of the Government Paperwork Elimination Act," 65 FR 25508, May 2, 2000 ("OMB Guidance"). As part of the OMB Guidance, the Department of Justice issued practical guidance on legal considerations related to agency use of electronic filing and recordkeeping. It should be noted that the E-SIGN Act ("Electronic Records and Signatures in Global and National Commerce Act")[58] is of importance in this matter.[59]

The GPEA, like the E-SIGN Act, provides that electronic records and signatures shall not be denied legal effect because they are in electronic form. Thus, a statutory "writing" requirement does not necessarily imply that this writing must be on paper.[60]

> "The shift away from paper-based records raises serious record collection, management and retention issues, some of which are familiar to the world of paper records and some of which are unique to electronic record retention and retrieval. On the other hand, electronic records can offer benefits, like easier search and retrieval, that may reduce some of the problems of paper-based records management. Thus, the objective of any conversion to electronic processes is to maximize the benefits that such systems can offer, while simultaneously minimizing any risks, including legal risks."[61]

> "A key question agencies face in converting to or adopting electronic processes is whether the system under consideration meets the applicable legal requirements and provides adequate evidence of its transactions and actions. In certain situations, an agency may determine that an electronic process is "good enough" to meet its legal

[58] Pub.L. 106-229, §1, June 30, 2000, 114 Stat. 464, codified at 15 U.S.C. 7001-7006.
[59] DOJ GPEA Guide, p. 1.
[60] DOJ GPEA Guide, p. 13.
[61] DOJ GPEA Guide, p. 2.

> needs without regard to whether it is comparable to or as good as its prior process. At the other extreme, some agencies may decide that electronic conversion will require a complete re-engineering of their business processes in order to address the legal risks and issues that a particular system presents or that are not being addressed as effectively in their existing system."[62]
>
> "To be able to protect the government's interests in litigation, the Department of Justice needs available, reliable, and persuasive agency records: records that are complete, uniform, easily understood, easily accessible and have been kept under a system that ensures a chain of custody of submissions and information gathered from all sources. Those requirements will not disappear merely because the medium of transactions changes from paper to electronic."[63]

As an agency identifies processes for conversion from paper to electronic, it should address the following issues:[64]

1. Will the electronically gathered and stored information be collected, retained, and accessible whenever needed?
2. Will the electronic collection, transmission, or storage of "documents" or information comply with applicable legal requirements, including, for example, laws requiring that certain records be maintained in a particular form or format?
3. Will electronic records be sufficiently reliable to be useful to Congress, agency decision-makers, private disputants, judges, juries, and others who must determine the facts underlying agency actions?
4. Will the agency's use of electronic methods to obtain, send, disclose and store information comply with applicable laws, such as those governing recordkeeping, privacy, confidentiality, and accessibility?

Will the electronic process gather all necessary information, meaning content, context, transaction data and identification data?

Content

The content of the transaction must include all records that comprise the substance of the transaction or filing.[65]

[62] DOJ GPEA Guide, p. 5.
[63] DOJ GPEA Guide, p. 7.
[64] DOJ GPEA Guide, p. 8.
[65] This was the subject of debate in Public Citizen v. Carlin, 184 F.3d 900, 910 (D.C. Cir. 1999) (discussing preservation of content, structure, and context of federal records).

> "When agencies collect information in paper form, additional information beyond that requested by the four corners of the form is frequently supplied. The document received by an agency might include additional attachments not necessarily required, and the agency might supplement the record with interlineations or notes. The physical composition of the document can attest to its completeness – for example, pages that were stapled together by the sender suggest that this was the document that was intended to be submitted to the agency. The agency's electronic process should include safeguards so that an agency can establish all of the information that was submitted by the sender as a single electronic document."[66]

Clearly, the distinction between final document and draft should be made,[67] and in general some form of version control is necessary.

Contextual information

When forms are used, not just the answers given by respondants but also the questions in the form should be preserved.[68] When information derives from electronic processes, sufficient documentation of the logic behind them and the guarantees in place to ensure reliability of their outcome should be available.[69]

The relationship between records must be preserved.[70]

Transaction data

Records that contain information about how the transaction was processed, including dates received and changes or modifications that were made in records should be preserved.

> "Agencies should also ensure that their electronic processing captures all relevant information, such as when and where the document was sent and received and whether the document was subsequently altered, and, if so, the source, date, and content of the alteration. Electronic systems can be designed to capture such information, including alterations or changes to a document. In the above example, if the agency's electronic process reliably kept track of all alterations to the applications after receipt, it could prove that Baker's application was not altered."[71]

[66] DOJ GPEA Guide, p. 10.
[67] DOJ GPEA Guide, p. 37.
[68] DOJ GPEA Guide, p. 21.
[69] DOJ GPEA Guide, p. 22 and 36.
[70] DOJ GPEA Guide, p. 37.
[71] DOJ GPEA Guide, p. 10.

The date and time that the communication or transaction was sent or initiated should be recorded,[72] as well as the fact that the communication or transmission actually was received, by whom it was received, and the date and time it was received.[73]

Identification

A means to authenticate the identity of all people who participated in the transaction both inside and outside the agency, as well as the scope of each person's participation should be included.

> "Often it is crucial to be able to prove who (i.e., a specific individual) submitted a communication or agreed to a transaction with an agency. Paper documents generally accomplish this fairly well, most commonly by containing a handwritten signature that can be matched with a specific person, a letterhead or return address on the document or envelope, and so on. Some transactions are so important that agencies require a personal appearance before some designated official in order to establish identity, e.g., having a notary endorse or certify the signature.
>
> Agencies should consider whether, in appropriate circumstances, a proposed electronic process will gather information sufficient to identify the person who submitted a communication or agreed to a transaction. For important transactions, particularly those that require proof of an individual's identity, or that he or she is creating a legally binding obligation, an agency may wish to require those individuals to employ some form of electronic signature. In the above example, the use of a digital signature could provide the agency a reliable means of identifying the name, position, and location of the specific individual who submitted the document, and thus, it would be difficult for Company to deny that one of its employees filed the application."[74]

Intention

For appropriate transactions, a means for establishing the intent of the participants to enter into the transaction or agreement should be provided.

> "Enforcement of an agency's rights often depends upon being able to prove what was intended by a communication. Did the parties intend a transmission of information to be a draft of a possible contract or a final, legally binding contract? Did

[72] DOJ GPEA Guide, p. 34.
[73] DOJ GPEA Guide, p. 36.
[74] DOJ GPEA Guide, p. 10 ff and 34 ff.

> an individual who transmitted information to the agency intend it to be a formal report which, if false, could result in his criminal prosecution? Paper-based transactions and communications typically answer such questions in a number of ways, for example, by whether a document "looks" like a contract or just an informal letter, whether it contains a handwritten signature, or whether it contains a warning that it is submitted under "penalty of perjury." Similar methods can be used in the electronic world."[75]

Will the information be retained?

Agencies should determine which information should be retained and for what period of time, as well as which information may be discarded soon after receipt.[76]

> "Electronic systems should be designed and maintained to guard against data corruption, whether through accidental deletion, equipment failures, storage media deterioration over time, stray electromagnetic forces, or myriad other hardware and software problems. Such systems should also be designed to limit access to authorized users – for example, by requiring controlled password identification for access to certain information. Finally, an electronic system should be designed to ensure proper file retention and tracing of alterations and updates (as to source, date, and content, and all other internal controls that are required to produce a secure and reliable record maintenance and retention system)."[77]

> "Electronic data are frequently transferred or converted from one storage medium or software system to another. In this process (sometimes referred to as "data migration"), important information, such as formatting and the structure and content of electronic forms, may be lost, or even the record itself destroyed unless appropriate steps are taken. Similarly, unless such changes are thoroughly documented, it can be difficult to demonstrate that the critical information was not changed in the process. In transition between systems, agencies sometimes maintain multiple, overlapping systems, particularly in the transition from paper to electronic based systems. Because information from all systems may be required to be maintained under the Federal Records Act[78] and may be needed for various purposes, agencies should address retention issues for all systems, even overlapping ones."[79]

[75] DOJ GPEA Guide, p. 11. See also p. 36.
[76] DOJ GPEA Guide, p. 11.
[77] DOJ GPEA Guide, p. 11 ff.
[78] 44 U.S.C. §§ 2101-2118, 2901-2910, 3101-3107, and 3301-3324.
[79] DOJ GPEA Guide, p. 12.

Not only must information be retained, it must be preserved under such conditions that guarantee it's integrity to a sufficient degree. Otherwise, the agency, judges or citizens may be unwilling to rely on the records concerned in the course of their activities.[80] This is of great importance to ensure the admissibility of the records as evidence, which depends on the evidence being shown to be authentic and to conform with the 'best evidence' rule.[81]

Will the information continue to be accessible?

> "Unlike paper files which, when properly organized and maintained in the ordinary course of business, are readily available and usable without any special equipment, electronic information is not always accessible without special equipment and software. Agencies should consider several factors related to the accessibility of electronic records. First, computer technology is rapidly changing and software and formatting standards may quickly become obsolete. Computer-stored data may become useless unless the agency can provide the continued capability with the older technologies or can accurately translate the document as more modern systems are implemented. Second, if in the future, an agency no longer has staff who are familiar and competent to work with the electronic processes necessary to read older data, such data could be functionally unavailable.[82] Electronic files might be stored while encrypted by software or protected by passwords no longer available or remembered years later, unless steps are taken to preserve the software or passwords. As noted above, these concerns are no less serious if the information is held by an outside party."[83]

For some specific guidances, see DOJ GPEA Guide, p. 40 ff.

What are the legal requirements?

Legal requirements generally pertain to creation[84], use, storage and disclosure.[85]

Does the electronic record constitute a 'writing' in situations where the law requires it?

The functional purposes of a writing must be attained by the electronic documents, meaning that they provide a documentary recording of a transaction in a manner that establishes and memorializes the terms.[86] In some instances, electronic information resembles an oral conversation

[80] DOJ GPEA Guide, p. 21 and 36 ff. The DOJ GPEA Guide uses the term 'perceived reliability' in this context.
[81] DOJ GPEA Guide, p. 22 **?**, .
[82] See Jeff Rothenberg, Ensuring the Longevity of Digital Documents, Scientific American, January 1995, at 42-47.
[83] DOJ GPEA Guide, p. 12.
[84] Notably, records retention obligations and requirements of form impact creation of electronic records.
[85] For a list of examples of these requirements, see: DOJ GPEA Guide, p. 23 ff.
[86] DOJ GPEA Guide, p. 13.

more than a formal agreement, e.g. chat conversations. Allthough such electronic conversations may be recorded and retained, judges may very well be reluctant to accord them the status and legal effect of a 'writing'."To the extent that the electronic process clearly records the terms of agreements and is adequate to show that the parties intended to make those agreements – that is, they serve the purposes that the law has required and relied on paper to serve – it is more likely that they will be accepted by the courts."[87]

5.3.2 Private sector

Recordkeeping in the private sector serves roughly 3 purposes:

- Compliance with legal requirements to maintain and preserve records
- Contractual obligations to maintain and preserve records
- Interest of the organisation in maintaining records as assets, ie for evidence purposes or for the purpose of retaining a memory of previous activities.

INT

The Basel II Accord modifies the Basel I accord on credit risk and extends it to address related issues.

Like Basel I, the accord determines rules on minimum capital requirements for financial institutions (first pillar of Basel II), though it no longer limits it's view to credit risk alone. Basel II also takes into account operational and market risks.

Operational risk encompasses the risk of loss due to operational issues (bad management decisions, fraud, incorrect administration, . . .). Operational risk is defined as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk.

Market risk is the risk of loss due to a decrease in the value of investments (e.g. stocks and bonds) held by the financial institution.

The second pillar of Basel II addresses the implementation of a risk management strategy by financial institutions throughout their activities. The aim is to encourage banks to develop and use better risk management techniques in monitoring and managing their risks.[88]

[87] DOJ GPEA Guide, p. 15, in particular footnote 19.
[88] Basel II, p. 204.

For risk management to be effective, a certain level of control over internal information an reporting must be ensured. The accuracy and completeness of data inputs into the bank's risk assessment process must be assured.[89]

The third pillar promotes transparancy by providing disclosre requirements, ie regarding the methodology for and results of risk calculations by financial institutions.

Further reading

- Bank for international settlements `http://www.bis.org`
- Basel II `http://www.bis.org/publ/bcbsca.htm`

EU

EU Directive on Statutory Audit

On July 7th 2006 the "8th EU Directive on Statutory Audit" was issued, replacing the EU 8th Company Law Directive of 1984. This directive must be transposed by the Member States by July 2008. The revision was first proposed in the Green paper on Financial Services Policy (2005-2010)[90]

This directive is the European counterpart of the U.S. Sarbanes-Oxley Act, which is why it is often referred to as Euro-SOX. The directive is less farreaching than its U.S. counterpart according to analists. While Sarbanes Oxley is based on rules, EuroSox is based on principles. EuroSox follows a 'comply or explain' approach to compliance demands.

The revision revolves around four key issues[91]:

1. establishing that board members are collectively responsible for financial statements and key non-financial information;
2. making unlisted companies' transactions with related parties more transparent;
3. ensuring that all companies provide full information about off-balance-sheet arrangements, including special-purpose vehicles which may be located offshore;
4. making listed companies issue an annual " corporate governance statement"

Revisions were made to the following directives:

[89] Basel II, p. 209, nr. 745.

[90] `http://ec.europa.eu/internal_market/finances/docs/actionplan/index/green_en.pdf`.

[91] Commission press release IP/04/1318 and MEMO/04/246, October 28, 2004.

- 4th directive 78/660/EEC – Annual Accounts of specific type of companies
- 7th directive 83/349/EEC – Consolidated accounts
- 8th directive 84/253/EEC – Auditor and audit committee requirements

In short, the requirements revolve around

- Effective Corporate Governance, internal controls and risk management
- Safeguard shareholders' investments
- Increase in disclosure requirements
- Establish audit committees
- Improved Corporate Governance.

These directives address a need for more reliable financial reporting, which is only one element in the larger frame of overal compliance with regulations.

US

Sarbanes-Oxley Act

Much like the FDA part 11 rules, the Sarbanes-Oxley Act – also known as the Public Company Accounting Reform and Investor Protection Act of 2002[92] and commonly called SOx – calls for (amongst other measures) internal controls over especially important company records, in this case records contributing to accurate financial reporting.

Key provisions of the act relating to records management are sections 302, 404 and 802.[93]

Section 302 of the SOx Act[94] requires that internal procedures be implemented to ensure accurate financial disclosure. The signing officers must certify that they are

- "responsible for establishing and maintaining internal controls"
- "have designed such internal controls to ensure that material information relating to the company and its consolidated subsidiaries is made known to such officers by others within those entities, particularly during the period in which the periodic reports are being prepared."
- "have evaluated the effectiveness of the company's internal controls as of a date within 90 days prior to the report"

[92] Pub.L. 107-204, 116 Stat. 745, enacted 30 July 2002.

[93] NEARON, BRUCE H. et al., "Life after Sarbanes-Oxley: The merger of information security and accountability", *Jurimetrics*, 45 2005, p. 380.

[94] Codified in 15 USC § 7241(a)(4).

- "have presented in the report their conclusions about the effectiveness of their internal controls based on their evaluation as of that date."

The SEC issued rules in 17 CFR §§240.13a-14 and 15d-15 pursuant this provision.

Section 404 of the SOx Act[95] requires management and the external auditor to report on the adequacy of the company's internal control over financial reporting.

More specific legal obligations pursuant this section have been issued by the SEC in "Final Rule: Management's Report on Internal Control Over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports" [96]

Furthermore, management can refer to the guidance on internal control reporting issued by the SEC in "Commission Guidance Regarding Management's Report on Internal Control Over Financial Reporting Under Section 13(a) or 15(d) of the Securities Exchange Act of 1934".[97]

External auditors may find guidance in Auditing Standard No. 5 of the Public Company Accounting Oversight Board (PCAOB).[98]

Though quite some regulatory texts have resulted from the SOx Act, critics find that the requirements remain all too vague.[99]

Section 802[100] provides that "Whoever knowingly alters, destroys, mutilates, conceals, covers up, falsifies, or makes a false entry in any record, document, or tangible object with the intent to impede, obstruct, or influence the investigation or proper administration of any matter within the

jurisdiction of any department or agency of the United States or any case filed under title 11, or in relation to or contemplation of any such matter or case, shall be fined under this title, imprisoned not more than 20 years, or both." Additionally, external auditors are required to retain corporate audit records for 5 years.[101]

Compliance with these provisions logically requires that an appropriate records management policy is in place.[102]

[95] Codified in 15 USC § 7262.

[96] Release No. 33-8238 (June 5,2003), available at `http://www.sec.gov/rules/final/33-8238.htm`.

[97] Release No. 33-8810 (June 20, 2007), available at `http://www.sec.gov/rules/interp/2007/33-8810.pdf`.

[98] `http://www.pcaob.org/Rules/Docket_021/2007-06-12_Release_No_2007-005A.pdf`.

[99] MONTANA, JOHN, "The Sarbanes-Oxley Act: Five Years Later", *The Information Management Journal*, 2007, Nr. Nov/Dec, p. 48.

[100] Codified in 18 USC §1519.

[101] See 18 USC §1520.

[102] Nearon et al. use the term 'Information Security Regime',which encompasses records management since "[information] provides the records and forensic histories for management, audit, compliance, records retention, and other functions", NEARON et al., *Jurimetrics 45 [2005]*, *op. cit.* (as in n. **??**) p. 381 and 382.

This follows directly from the definition of 'internal control' given by the SEC in it's final rules[103]:

- 17 CFR §240.13a15 (f) The term **internal control** over financial reporting is defined as a process designed by, or under the supervision of, the issuer's principal executive and principal financial officers, or persons performing similar functions, and effected by the issuer's board of directors, management and other personnel, to provide reasonable assurance regarding the **reliability** of financial reporting and the preparation of financial statements for external purposes in accordance with generally accepted accounting principles and includes those policies and procedures that:
  1. Pertain to the **maintenance of records** that in reasonable detail accurately and fairly reflect the transactions and dispositions of the assets of the issuer;
  2. Provide reasonable assurance that **transactions are recorded** as necessary to permit preparation of financial statements in accordance with generally accepted accounting principles, and that receipts and expenditures of the issuer are being made only in accordance with authorizations of management and directors of the issuer; and
  3. Provide reasonable assurance regarding prevention or timely detection of unauthorized acquisition, use or disposition of the issuer's assets that could have a material effect on the financial statements.

In their article, Nearon et al. introduce a set of specific terms, in particular 'data-generating events' and 'source data'.

Though the definition they give of 'data-generating event' does not correspond entirely with that of a 'record', the concerns they raise with respect to DGE's apply to records. Main concerns are tampering with the time and date of records, as well as tampering with the contents.[104] A records management system must ensure integrity, confidentiality (as appropriate) and availability (readability).[105] A valid point raised by Nearon et al. is that the (human readable) views generated on the basis of electronically stored data must be reliable as well. Indeed there is little point in securely storing records if the software used to serve them to the user can be manipulated to show something different.[106]

The annually required report on the company's internal control over financial reporting must include a statement "identifying the framework used by management to evaluate the effective-

[103] "Final Rule: Management's Report on Internal Control Over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports", Release No. 33-8238 (June 5, 2003), available at `http://www.sec.gov/rules/final/33-8238.htm`.

[104] NEARON et al., *Jurimetrics 45 [2005]*, *op. cit.* (as in n. **??**) p. 390 and 394.

[105] NEARON et al., *Jurimetrics 45 [2005]*, *op. cit.* (as in n. **??**) p. 394.

[106] NEARON et al., *Jurimetrics 45 [2005]*, *op. cit.* (as in n. **??**), p. 387 ff.

ness of the company's internal control over financial reporting."[107] The framework used must be "a suitable, recognized control framework that is established by a body or group that has followed due-process procedures, including the broad distribution of the framework for public comment."[108] The SEC acknowledges that there are many different ways to conduct an evaluation of the effectiveness of internal control over financial reporting, to aid companies is selecting a method it has issued an interpretive guidance document.[109]

The only industry standard explicitly mentioned is COSO[110] Internal Control – Integrated Framework[111], wich offers very little specific guidance on information security.[112]

The IT Governance Institute has explored the significance of SOx with respect to information technology.[113] Records management is not addressed in any significant detail by this report.

In the U.S. the rules regarding discovery pose particular requirements for preservation of documents. The impact of these rules is on electronic records management is the source of numerous disputes.

The Sedona Conference Institute[114] has a working group on Electronic Document Retention and Production [115], which has published The Sedona Principles and The Sedona Guidelines on this topic.

The Sedona Guidelines state that:

1. An organization should have reasonable policies and procedures for managing its information and records.
2. An organization's information and records management policies and procedures should be realistic, practical, and tailored to the circumstances of the organization.
3. An organization need not retain all electronic information ever generated or received.
4. An organization adopting an information and records management policy should also develop procedures that address the creation, identification, retention, retrieval, and ultimate disposition or destruction of information and records.

[107] Final Rule: Management's Report on Internal Control Over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports, Release No. 33-8238 (June 5, 2003), available at `http://www.sec.gov/rules/final/33-8238.htm`.

[108] 17 CFR §240.13a-15 and 15d-15 (c).

[109] Release No. 34-55929.

[110] Committee of Sponsoring Organizations of the Treadway Commission, `http://www.coso.org`.

[111] See the background information in the Final Rule: Management's Report on Internal Control Over Financial Reporting and Certification of Disclosure in Exchange Act Periodic Reports, Release No. 33-823.

[112] NEARON et al., *Jurimetrics 45 [2005]*, *op. cit.* (as in n. **??**), p. 409 ff.

[113] IT Governance Institute, *IT Control Objectives for Sarbanes-Oxley*, 2004, 34 p.

[114] `http://www.thesedonaconference.org/`.

[115] `http://www.thesedonaconference.org/wgs`.

5. An organization's policies and procedures must mandate the suspension of ordinary destruction practices and procedures as necessary to comply with the preservation obligations related to actual and reasonably anticipated litigation, government investigation, or audit.

Further reading

- Sedona Conference Institute `http://www.thesedonaconference.org/`
- MASON, STEPHEN, "Authentic Digital Records: Laying the Foundation for Evidence", *The Information Management Journal*, 2007, Nr. Sept/Okt
- JONES, THOMAS M. et al., "Going global. Mapping an international records retention strategy." *The Information Management Journal*, 2008, Nr. May/June

Bibliography

Agnoloni, Tommaso, Francesconi, Enrico and Spinosa, Pierluigi, "xmLegesEditor: an Open-Source Visual XML Editor for supporting Legal National Standards", In **Biagioli, Carlo, Francesconi, Enrico and Sartor, Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art17.pdf, 239–251.

Ahmad, A., "The Forensic Chain of Evidence Model: Improving the Process of Evidence Collection in Incident Handling Procedures", In **X. (ed.)**, *Proceedings of the 6th Pacific Asia Conference on Information Systems*, 2002 , http://www.dis.unimelb.edu.au/staff/atif/AhmadPACIS.pdf, 5 p.

Aiken, Peter et al., *Measuring Data Management Practice Maturity: A Community's Self-Assessment*, 2007 , http://doi.ieeecomputersociety.org/10.1109/MC.2007.139.

Ajani, Gianmaria et al., "Multilingual Conceptual Dictionaries Based on Ontologies", In **Biagioli, Carlo, Francesconi Enrico Sartor Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art12.pdf, 161–172.

Anderson, Anne, *A Comparison of Two Privacy Policy Languages:EPAL and XACML*, Sun Microsystems Laboratories, 2005 , http://research.sun.com/techrep/2005/smli_tr-2005-147/TRCompareEPALandXACML.html, Technical Report.

Bearman, David et al., "A common model to support interoperable metadata", *D-Lib Magazine*, Vol. 5 1999, Nr. 1 , http://www.dlib.org/dlib/january99/bearman/01bearman.html.

Becker, Arnd, *Elektronische Dokumente als Beweismittel im Zivilprozess*, Frankfurt, Peter Lang, 2004.

Berger, Christian, "Beweisführung mit elektronischen Dokumenten", *NJW*, 2005, Nr. 15, 1016–1020.

Biagioli, C. et al., "The NIR Project: standards and tools for the Italian legislative environment", Berlin, 2004 , http://www.jurix.nl/index.php?option=com_docman\task=docclick\Itemid=27\bid=14\limitstart=0\limit=10, Presentation.

Biagioli, Carlo, "How to link (external) models or interpretations of the meaning of sources of law to the original sources", Leiden, 2007 , http://www.lri.jur.uva.nl/~winkels/PP-Jurix-2007.pdf.

Biagioli, Carlo et al., "Law Making Environment. Perspectives", In **Biagioli, Carlo, Francesconi, Enrico and Sartor, Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art19.pdf, 267–281.

Binazzi, Simona et al., "ITLaw: An Advanced Documentation System in Legal Informatics", *The Journal of Information, Law and Technology*, 1999, Nr. 1 , http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/1999_1/idg/binazzi/.

Blarkom, G.W. van, Borking, J.J. and Olk, J.G.E., *Handbook of Privacy and Privacy-Enhancing Technologies, The case of Intelligent Software Agents*, The Hague, College bescherming persoonsgegevens, 2003 , http://www.andrewpatrick.ca/pisa/handbook/Handbook_Privacy_and_PET_final.pdf.

Boer, Alexander, "Using event descriptions for metadata about legal documents", In **Winkels, Radboud and Francesconi, Enrico (ed.)**, *Electronic Proceedings of the Workshop on Standards for Legislative XML*, 2007 , http://www.leibnizcenter.org/~winkels/events.pdf, in conjunction with Jurix 2007.

Boer, Alexander, Hoekstra, Rinke and Winkels, Radboud, "Metalex: Legislation in XML", In **Bench-Capon, Trevor, Daskalopulu, Aspassia and Winkels, Radboud (ed.)**, *Legal Knowledge and Information Systems: JURIX 2002*, IOS Press, 2002 , http://www.jurix.nl/pdf/j02-01.pdf.

Borking, J.J., "The status of Privacy Enhancing Technologies", In **Nardelli, E., Posadziejewski, S. and Talamo, M. (ed.)**, *Certification and Security in E-Services, From E-Government to E-Business*, Boston, Kluwer, 2003, 211–246.

Borking, J.J. et al., *Methodology of Privacy Threat Analysis*, The Hague, EU PISA project IST-2000-26038, 2001, Deliverable 7 of WP 2.

Borking, John, "Privacy Rules, A Steeple Chase For Systems Architects", In **X. (ed.)**, *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, Ispra, Italy, W3C, 2006 , http://www.w3.org/2006/07/privacy-ws/papers/, 17 p.

Broucek, Vlasti and Turner, Paul, "Winning the Battles, Losing the War? Rethinking Methodology for Forensic Computing Research", *Journal in Computer Virology*, Vol. 2 2006, Nr. 1, 3–12.

Broucek, Vlasti, Turner, Paul and Frings, Sandra, *Music piracy, universities and the Australian Federal Court: Issues for forensic computing specialists*, 2005, 21 , http://dx.doi.org/10.1016/j.clsr.2005.01.014, 30–37.

Bundesarchiv, Schweizerisches (ed.), *Rechtliche Aspekte elektronischer Geschäftsverwaltung*, Bern, 2007 , http://www.isb.admin.ch/themen/architektur/00078/00080/00212/index.html?lang=de\download=NHzLpZeg7t,, 17 p.

Casanovas, Pompeu et al. (ed.), Stanford, CA, USA,, Stanford University, 2007, http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-321/.

Casassa Mont, Marco, "Dealing with Privacy Obligations: Important Aspects and Technical Approaches", In **Katsikas, Sokratis K., Lopez Javier Pernul Günther (ed.)**, *Trust and Privacy in Digital Business*, Volume 3184, Lecture Notes in Computer Science, Springer, 2004 , http://dx.doi.org/10.1007/b99832, 120–131.

Casassa Mont, Marco, *A System to Handle Privacy Obligations in Enterprsies*, HP, 2005, HP Technical Report , http://www.hpl.hp.com/techreports/2005/HPL-2005-180.html, 104 p, HPL-2005-1.

Casassa Mont, Marco; X. (ed.), *On the Need to Explicitly Manage Privacy Obligation Policies as Part of Good Data Handling Practices*, Ispra, Italy, W3C, 2006 , http://www.w3.org/2006/07/privacy-ws/papers/, 4 p.

Cevenini, Claudia et al., "Development of the ALIS IP Ontology: Merging Legal and Technical Perspectives", In **X. (ed.)**, *Computer-Aided Innovation (CAI), IFIP International Federation for Information Processing*, Boston, Springer, 2008 , http://dx.doi.org/10.1007/978-0-387-09697-1_14, 169–180.

Committee, PREMIS Editorial, *PREMIS Data Dictionary for Preservation Metadata*, PREMIS, 2008 , http://www.loc.gov/standards/premis/, 217 p.

Coyle, Karen, "Descriptive metadata for copyright status", *First Monday*, 10 2005, Nr. 10 , http://www.firstmonday.org/issues/issue10_10/coyle/index.html.

Coyle, Karen, *Rights in the PREMIS Data Model*, Washington D.C., U.S.A., Library of Congress, 2006 , http://www.loc.gov/standards/premis/Rights-in-the-PREMIS-Data-Model.pdf, 32 p.

CTOSE Consortium, *CTOSE Project Results*, 2003 , http://web.archive.org/web/*hh_/www.ctose.org/ResultsPaperv6.pdf.

de Oliveira Lima, João Alberto, "An Adaptation of the FRBR Model to Legal Norms", In **Biagioli, Carlo, Francesconi Enrico Sartor Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art4.pdf, 53–65.

Dekeyser, Hannelore and Lipinski, Tomas, "Digital Archiving and Copyright Law: A Comparative Analysis", *International Journal of Communication Law and Policy*, 12 2008, 179–224 , http://www.ijclp.net/12_2008/pdf/dekeyserlipinski.pdf.

Dekkers, Makx, Weibel Stuart, "State of the Dublin Core Metadata Initiative", *D-Lib Magazine*, Vol. 9 2003, Nr. 4 , http://dlib.org/dlib/april03/weibel/04weibel.html.

Delgado, Jaime et al., "IPROnto - Intellectual Property Rights Ontology", *ISWC*, 2002 , http://dmag.upf.edu/ontologies/ipronto/ISWCPoster.pdf, Poster.

Delgado, Jaime et al., "IPROnto: An Ontology for Digital Rights Management", In **Bouncier, D. (ed.)**, *Legal Knowledge and Information Systems*, Jurix 2003, Amsterdam, IOS Press, 2003 , http://www.jurix.nl/, 111–120.

Dinant, Jean-Marc, "The long way from electronic traces to electronic evidence", *International Review of Law, Computers Technology*, Vol. 18 2004, 173–183.

EFD, Eidgenössisches Finanzdepartement (ed.), *Vorgaben zu GEVER / Records Management*, 2007 , http://www.isb.admin.ch/themen/architektur/00078/00197/00214/index.html?lang=de\download=NHzLpZeg7t,, 4 p.

ENFSI (ed.), *Guidelines for Best Practice in the Forensic Examination of Digital Technology*, ENFSI, 2006 , http://www.enfsi.eu/uploads/files/ENFSI_Forensic_IT_Best_Practice_GUIDE_5\%5B1\%5D.0.pdf, 28 p, Version 5.

Europe, RAND (ed.), *Handbook of Legislative Procedures of Computer and Network Misuse in EU Countries – Study for the European Commission*, Directorate-General Information Society, 2002, 287 p.

Fellbaum, Christiane, *WordNet: An electronic lexical database*, Cambridge, Mass., MIT Press, 1998, 305 p.

Finke, Nicholas D., "TEI Extensions for Legal Text", In **X. (ed.)**, *Proceedings of the Text Encoding Initiative Tenth Anniversary User Conference*, 1997 , `http://xml.coverpages.org/finkeTEI10.html`.

Fischer, Laurie, "Condition critical: developing records retention schedules", *The Information Management Journal*, Jan/Feb 2006, 26.

Francesconi, Enrico, "The "Norme in Rete"- project: Standards and tools for Italian legislation", *International Journal of Legal Information*, Vol. 34 2006, Nr. 2, 358–376 , `http://www.xmleges.org/ita/images/stories/francesconiijli06.pdf`.

Franks, Pat and Kunde, Nancy, "Why metadata matters", *The Information Management Journal*, Sept/Oct 2006, 55–61.

Frings, S. et al., "Cyber Crime Advisory Tool - C*CAT: a holistic approach to electronic evidence processing", *Proceedings of the 10th International Conference on Human-Computer Interaction*, 3 2003, 704–708.

García, Roberto, *A Semantic Web Approach to Digital Rights Management*, Barcelona, Spain, Department of Technologies, Universitat Pompeu Fabra, 2005 , `http://rhizomik.net/~roberto/thesis/`, 286 p.

García, Roberto, Gil, Rosa and Delgado, Jaime, "Intellectual Property Rights Management using a Semantic Web Information System", In **X. (ed.)**, *OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2004*, Lecture Notes in Computer Science, Berlin, Springer, 2004, 3291, 689 – 704.

García, Roberto, Gil, Rosa and Delgado, Jaime, "A web ontologies framework for digital rights management", *Artificial Intelligence and Law*, Vol. 15 2007, Nr. 2, 137–154 , `http://dx.doi.org/10.1007/s10506-007-9032-6`.

Gasser, Urs and Haeusermann, Daniel M., "E-Compliance: Konzept, Merkmale, Aufgaben und organisatorische Auswirkungen", In **X. (ed.)**, *Internet-Recht und Electronic Commerce Law: 9. Tagungsband*, Bern, Stämpfli, 2006, p. 71–100.

Gasser, Urs and Haeusermann, Daniel M., *E-Compliance: Towards a Roadmap for Effective Risk Management*, Harvard, The Berkman Center for Internet Society, 2007 , http://ssrn.com/abstract=971848, 24 p.

Giblin, C. et al., "Regulations Expressed as Logical Models (REALM)", In **Moens, Marie-Francine and Spyns, Peter (ed.)**, *Proceedings of the 18th Annual Conference on Legal Knowledge and Information Systems*, Brussels, Jurix, 2005, 37–48.

Gounaris, Anastasios and Theodoulidis, Babis, "Data Base Management Systems (DBMSs): Meeting the requirements of the EU data protection legislation", *International Journal of Information Management*, 23 2003, Nr. 3, 185–199 , http://dx.doi.org/10.1016/S0268-4012(03)00023-9.

Governatori, Guido and Rotolo, Antonino, "Modelling Contracts Using RuleML", In **Gordon, T. (ed.)**, *Legal Knowledge and Information Systems*, Jurix 2004, Amsterdam, IOS Press, 2004 , http://www.jurix.nl/pdf/j04-16.pdf, 141–150.

Grandi, Fabio, Mandreoli, Federica and Tiberio, Paolo, "Temporal modelling and management of normative documents in XML format", *Data Knowledge Engineering*, vol. 54 2005, Nr. 3, 327 – 354 , http://dx.doi.org/10.1016/j.datak.2004.11.002.

Group, Rights Data Management (ed.), *CopyrightMD User Guidelines, Version 0.9*, California Digital Library, 2006 , http://www.cdlib.org/inside/projects/rights/schema/copyrightMD_user_guidelines.pdf, 19 p.

Gunter, Carl A., "Ensuring Privacy Conformance in Inter-Domain Systems", In **X. (ed.)**, *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, Ispra, Italy, W3C, 2006 , http://www.w3.org/2006/07/privacy-ws/papers/, 5 p.

Hajtnik, Tatjana, "Maintaining legal value of a record throughout their lifecycle", *DLM Forum Meeting, Ljubljana, 8-9 april 2008*, 2008 , http://dlmforum.typepad.com/Slovenia_Hajtnik.pdf, Presentation.

Hatter, Clyde, "Standard Models for Legislation - The Cost of Compliance", In **Biagioli, Carlo, Francesconi, Enrico and Sartor, Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art16.pdf, 225–237.

Hietanen, Aki, "Networking European Legal Sites : Experiences and Challenges", In **X. (ed.)**, *Proceedings of the Law via the Internet Conference*, Paris, 2004 , http://www.frlii.org/IMG/pdf/hietanenparis.pdf.

Hirtle, Peter, *Copyright Term and the Public Domain in the United States 1 January 2007*, 2007 , http://www.copyright.cornell.edu/training/Hirtle_Public_Domain.htm, 4.

Hoekstra, Rinke et al., "The LKIF Core Ontology of Basic Legal Concepts", In **Casanovas, Pompeu et al. (ed.)**, *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*, Stanford, CA, USA,, 2007 , http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-321/.

Hoffmann, Mathis, "Der Beweiswert elektronischer Dokumente", *DSWR*, 2006, Nr. 3, 60 ff..

Hofman, Hans, "Developments in ISO standards for recordkeeping", 2005 , http://dlmforum.typepad.com/Paper_HansHofman_onstandards.pdf.

Hsu, Windsor W. and Ong, Shauchi, *Fossilization: A Process for Establishing Truly Trustworthy Records*, IBM Almaden Research Center, 2004, IBM Research , http://domino.research.ibm.com/library/cyberdig.nsf/1e4115aea78b6e7c85256b360066f0d4/02da1cea05c6c61, 11 p.

Iacovino, Livia and Todd, Malcolm, "The long-term preservation of identifiable personal data: a comparative archival perspective on privacy regulatory models in the European Union, Australia, Canada and the United States", *Archival Science*, vol. 7 2007, Nr. 1, 107–127 , http://dx.doi.org/10.1007/s10502-007-9055-5.

Iannella, Renato, "Digital Rights Management (DRM) Architectures", *D-Lib Magazine*, Vol. 7 2001, Nr. 6 , http://www.dlib.org/dlib/june01/iannella/06iannella.html.

Iannella, Renato, Henricksen, Karen and Robinson, Ricky, "A Policy Oriented Architecture for the Web: New Infrastructure and New Opportunities", Ispra, Italy, W3C, 2006 , http://www.w3.org/2006/07/privacy-ws/papers/05-ianella-policy-oriented-architecture, 4 p.

IFLA Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records*, München, K.G. Saur, 1998, 136 p.

Jones, Thomas M. et al., "Going global. Mapping an international records retention strategy." *The Information Management Journal*, 2008, Nr. May/June, 30–36.

Kabilan, Vandana, Johannesson, Paul and Rugaimukamu, Dickson M., "Business Contract Obligation Monitoring through Use of Multi Tier Contract Ontology", In **X. (ed.)**, *On The Move to Meaningful Internet Systems 2003: OTM 2003 Workshops*, Volume 2889, Lecture Notes in Computer Science, Berlin, Springer, 2003 , http://dx.doi.org/10.1007/b94345, 690–702.

Karyda, Maria and Mitrou, Lilian, "Internet forensics: legal and technical issues", *Proceedings of the second International Workshop on Digital Forensics and Incident Analysis* 2007.

Kearsley, Amanda J., "Electronic Document Management, Legal admissibility of evidence held in electronic form", *Computer Law Security Report*, Vol. 15 1999, Nr. 3, 185–187.

Keneally, Erin E., "Digital logs - proof matters", *Digital Investigation*, 2004, Nr. 1, 94–101 , http://www.elsevier.com/locate/diin.

Kenny, S. and Borking, J., "The Value of Privacy Engineering", *The Journal of Information, Law and Technology*, 2002, Nr. 1 , http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2002_1/kenny/.

Kerrigan, Shawn and Law, Kincho H., "Logic-based regulation compliance-assistance", New York, ACM Press, 2003, 126–135.

Klein, Susanne, "Die Beweiskraft elektronischer Verträge. Zur Entwicklung der zivilprozessrechtlichen Vorschriften über die Beweiskraft elektronischer Dokumente", *JurPC Web-Dok.* 2007, Nr. 198, 1–71.

Kuehl Frostestad, Heidi, "Globalex: A Unique and Valuable Tool for Foreign, Comparative, and International Law Research", *International Journal of Legal Information*, 34 2006, 473–482 , http://www.heinonline.org/HOL/Page?handle=hein.journals/ijli34\id=1\size=2\collection=journals\index.

Lavoie, Brian F., "PREMIS With a Fresh Coat of Paint. Highlights from the Revision of the PREMIS Data Dictionary for Preservation Metadata", *D-Lib Magazine*, Vol. 14 2008, Nr. 5/6 , http://www.dlib.org/dlib/may08/lavoie/05lavoie.html.

Lehmann, Jos et al. (ed.), *LOAIT - Legal Ontologies and Artificial Intelligence Techniques*, Volume 4, IAAIL Workshop Series, Tilburg, Wolf Legal Publishers, 2005.

Leroux, Olivier, "Legal admissibility of electronic evidence", *International Review of Law, Computers Technology*, Vol. 18 2004, Nr. 2, 193–220.

Limone, D. A., "L'insegnamento dell'informatica giuridica in Italia", In **Frosini, V and Limone, D. A. (ed.)**, *L'insegnamente dell'informatica giuridica*, Naples, 1990, p. 19–27.

Lyytikäinen, Virpi, Tiitinen, Pasi T. and Salminen, Airi, "XML Metadata for Accessing Heterogeneous Legal Databases", In **X. (ed.)**, *Proceedings of the XML Europe 2001 Conference*, 2001 , http://www.gca.org/papers/xmleurope2001/papers/html/s27-4.html.

Madsen, Paul, Casassa Mont, Marco and Wilton, Robin, "A Privacy Policy Framework - A Position paper for the W3C Workshop of Privacy Policy Negotiation", In **X. (ed.)**, *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, Ispra, Italy, W3C, 2006 , http://www.w3.org/2006/07/privacy-ws/papers/, 7 p.

Magnusson Sjöberg, Cecilia, Stockholm, Jure, 1998.

Marcus, J. Scott, Carter, Kenneth and Robinson, Neil, e.a., *Comparison of Privacy and Trust Policies in the Area of Electronic Communications*, Bad Honnef, wik-Consult GmbH, 2007 , http://ssrn.com/abstract=1086929, 214 p.

Mason, Stephen, "Authentic Digital Records: Laying the Foundation for Evidence", *The Information Management Journal*, 2007, Nr. Sept/Okt, 32–40.

Mason, Stephen, "Archiving and storing e-mails - The legal and practical issues", *Computer Law Security Report*, vol. 24 2008, Nr. 2, 176–180 , http://dx.doi.org/10.1016/j.clsr.2007.09.004.

May, Michael J., Gunter, Carl A. and Insup, Lee, "Privacy APIs: Access Control Techniques to Analyze and Verify Legal Privacy Policies", In **X. (ed.)**, *Computer Security Foundations Workshop*, Venice, Italy, 2006 , http://seclab.uiuc.edu/pubs/MayGL06.pdf, 13 p.

Miller, George A., "WordNet: A lexical database for english", *Communications of the ACM*, vol. 38 1995, Nr. 11, 39–41.

Mitrakas, Anreas, Zaitch Damien, "Law, Cybercrime and digital forensics: Trailing Digital Suspects", In **Kanelis, Panagiotis, Kiountouzis Evangelos Kolokotronis Nicholas Drak-**

oulis Martakos (ed.), *Digital Crime and Forensic Science in Cyberspace*, London, Idea Group, 2006, 267–290.

Moens, Marie-Francine, "Innovative techniques for legal text retrieval", *Artificial Intelligence and Law*, 9 2001, 29–57.

Montana, John, "The Sarbanes-Oxley Act: Five Years Later", *The Information Management Journal*, 2007, Nr. Nov/Dec, 48–53.

Muller, Murk, "Legal RDF Dictionary", In **X. (ed.)**, *Proceedings of XML Europe 2002*, 2002 , `http://www.idealliance.org/papers/xmle02/dx_xmle02/papers/03-04-03/03-04-03.html`.

Myler, Ellie, "Minimizing Risks through a corporate information compliance initiative", *The Information Management Journal*, 2008, Nr. Jan/Feb, 58–63.

Nearon, Bruce H. et al., "Life after Sarbanes-Oxley: The merger of information security and accountability", *Jurimetrics*, 45 2005, 379–412.

Palmirani, Monica and Brighi, Raffaella, "Time Model for Managing the Dynamic of Normative System", In **Wimmer, Maria A. et al. (ed.)**, *Electronic Government*,.

Peglar, Rob, "Evidence management solutions for mitigating e-records risks", *The Information Management Journal*, 2007, Nr. July/Aug, 56–60.

Perez Asinari, Maria Veronica, "Legal constraints for the protection of privacy and personal data in electronic evidence handling", *International Review of Law, Computers Technology*, Vol. 18 2004, Nr. 2, 231–250.

Peters, Wim, Sagri, Maria-Teresa and Tiscornia, Daniela, "The structuring of legal knowledge in LOIS", *Artificial Intelligence and Law*, Vol. 15 2007, Nr. 2, 117–135.

Petersen, Knud Erik, "Lex Dania XMl status april 2005", *I Quaderni*, 2005, Nr. 18, 13–19 , `http://www.cnipa.gov.it/site/_files/Quaderno\%2018.pdf`.

Pfitzmann, Birgit, Powers, Calvin and Waidner, Michael, *IBM's Unified Governance Framework (UGF) Initiative*, Zurich, IBM Research Division, 2007, IBM Research Report RZ 3699 , `http://www.zurich.ibm.com/pdf/csc/rz3699_UGF-whitepaper.pdf`, 22 p.

Riveret, Régis, Palmirani, Monica and Rotolo, Antonino, "Legal Consolidation formalised in Defeasible Logic and based on Agents", In **Biagioli, Carlo, Francesconi Enrico Sartor Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic

Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art9.pdf, 117–135.

Roßnagel, Alexander, Pfitzmann Andreas, "Der Beweiswert von E-Mail", *NJW*, 2003, Nr. 17, 1209–1214.

Rowlingson, Robert, "A Ten Step Process for Forensic Readiness", *International Journal of Digital Evidence*, vol. 2 2004, Nr. 3 , http://www.utica.edu/academic/institutes/ecii/publications/articles/A0B13342-B4E0-1F6A-156F501C49CF5.

Rubino, Rossella, Rotolo, Antonino and Sartor, Giovanni, "An OWL Ontology of Norms and Normative Judgements", In **Biagioli, Carlo, Francesconi, Enrico and Sartor, Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art13.pdf, 173–187.

Rundle, Mary, "International Personal Data Protections and Digital Identity Management Tools", In **X. (ed.)**, *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, W3C, 2006 , http://www.w3.org/2006/07/privacy-ws/papers/21-rundle-data-protection-and-idm-tools, 4 p.

Rust, Godfrey, Bide Mark (ed.), *The indecs metadata framework, Principles, model and data dictionary*, Indecs, 2000 , http://www.doi.org/topics/indecs/indecs_framework_2000.pdf, 49 p.

Rust, Godfrey, "Metadata: The Right Approach", *D-Lib Magazine*, 1998 , http://www.dlib.org/dlib/july98/rust/07rust.html.

Ryan, Henry et al., "Ontology-Based Platform for Trusted Regulatory Compliance Services", In **X. (ed.)**, *On The Move to Meaningful Internet Systems 2003: OTM 2003 Workshops*, Volume 2889, Lecture Notes in Computer Science, Berlin, Springer, 2003 , http://dx.doi.org/10.1007/b94345, 675–689.

Sato, O, Broucek, V and Turner, P., "Electronic evidence management for computer incident investigations: a prospect of CTOSE", *Security Manage*, 2005, Nr. 18, 11–18.

Sommer, P., "Intrusion Detection Systems as Evidence", In **X. (ed.)**, *Proceedings of Recent Advances in Intrusion Detection 1998*, 1998 , http://www.raid-symposium.org/raid98/Prog_RAID98/Full_Papers/Sommer_text.pdf, 14 p.

Spinosa, Pierluigi, "Expansion and Internationalization of the Italian Schema of Assignment of Uniform Names", *I Quaderni*, 2005, Nr. 18, 118–133 , http://www.cnipa.gov.it/site/_files/Quaderno\%2018.pdf.

Spinosa, Pierluigi, "Internationalization of the Legal URN Schema", In **Biagioli, Carlo, Francesconi, Enrico and Sartor, Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art6.pdf, 87–97.

Stephenson, P., "End-to-End Digital Forensics", *Computer Fraud and Security*, 2002, Nr. 9, 17–19 , http://dx.doi.org/10.1016/S1361-3723(02)00914-4.

Stephenson, P., "A Comprehensive Approach to Digital Incident Investigation", *Information Security Technical Report*, Vol. 8 2003, Nr. 2, 42–54 , http://dx.doi.org/10.1016/S1363-4127(03)00206-1.

Stephenson, P., "Using Evidence Effectively", *Computer Fraud and Security*, 2003, Nr. 3, 17–19 , http://dx.doi.org/10.1016/S1361-3723(03)03012-4.

Tillett, Barbara, *What is FRBR? A conceptual model for the bibliographic universe*, Washington D.C., U.S.A., Library of Congress, 2004 , http://www.loc.gov/cds/downloads/FRBR.PDF.

Tiscornia, Daniela, "Metadata for Content Description", *I Quaderni*, 2005, Nr. 18, 134–144 , http://www.cnipa.gov.it/site/_files/Quaderno\%2018.pdf.

Tiscornia, Daniela, "The Lois Project: Lexical Ontologies for Legal Information Sharing", In **Biagioli, Carlo, Francesconi, Enrico and Sartor, Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art14.pdf, 189–204.

Turtle, H., "Text retrieval in the legal world", *Artificial Intelligence and Law*, 3 1995, 5–24.

Van Engers, T. et al., "POWER: Using UML/OCL for Modeling Legislation -an application report", In **X. (ed.)**, *Proceedings of the 8th International Conference on Artificial Intelligence and Law (ICAIL 2001)*, New York, ACM, 2001, 157–167.

Vitali, Fabio and Zeni, Flavio, "Towards a country-independent data format: the Akoma Ntoso experience", In **Biagioli, Carlo, Francesconi, Enrico and Sartor, Giovanni (ed.)**, *Proceedings of of the V Legislative XML Workshop*, European Press Academic Publishing, 2007 , http://www.e-p-a-p.com/dlib/9788883980466/art5.pdf, 67–86.

Vossen, Piek, *EuroWordNet A Multilingual Database with Lexical Semantic Networks*, Dordrecht, Kluwer Academic Publishers, 1998, 184 p.

Waldron, Martin, "Adopting electronic records management: European strategic initiatives", *The Information Management Journal*, 2004, Nr. July/Aug, 30–35.

Weitzner, Daniel J. et al., "Transparency and End-to-End Accountability: Requirements for Web Privacy Policy Languages", In **X. (ed.)**, *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006*, W3C, 2006 , `http://www.w3.org/2006/07/privacy-ws/papers/`, 5 p.

Winkels, R.G.F. et al., "Generating Exception Structures for Legal Information Serving", In **Gordon, Th.F. (ed.)**, *Proceedings of the Seventh International Conference on Artificial Intelligence and Law (ICAIL-99)*, New York, ACM, 1999, 182–195.

Wolfe, Henry B., "Evidence Analysis", *Computers Security*, Vol. 22 2003, Nr. 4, 289–291 , `http://dx.doi.org/10.1016/S0167-4048(03)00404-8`.

X. (ed.), *The admissibility of electronic evidence in court: fighting against high-tech crime*, Barcelona, Cybex, 2006 , `http://www.cybex.es/agis2005/docs/libro_aeec_en.pdf`, 64 p.

Zhu, Qingbo, Hsu Windsor W., "Fossilized Index: The Linchpin of Trustworthy Non-Alterable Electronic Records", In **X. (ed.)**, *International Conference on Management of Data archive. Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, New York, ACM Press, 2005 , `http://portal.acm.org/citation.cfm?id=1066157.1066203`, 395–406.