Developments in ISO standards for recordkeeping

Hans Hofman, Nationaal Archief Netherlands

Introduction

Official standards for recordkeeping only emerged quite recently, but already have a big impact. The responsible committee, ISO TC46/SC11 records management, started in 1997 and attracts a growing number of members. Currently, there are 25 voting and about 15 observing members.[1]
Within the committee several working groups are covering different aspects of recordkeeping, including metadata, functional requirements for long term preservation, work process analysis, and access. The most well-known published standard is the ISO 15489:2001 on records management. In 2004 the review of this standard has started.
With respect to metadata for records a Technical Specification has been published in Spring 2004, ISO 23081-1. Other standards or technical reports with respect to records management are under way.
The work in ISO TC46/SC11 reflects a growing professionalism within records and archives management and what is seen as a common understanding at this moment of recordkeeping. As such it also tries to include and synthesise developments in different initiatives around the world. The results of projects such as Inter Pares, the Australian RKMS project, or the Clever Metadata project of Monash University in Melbourne for example feed into the ISO work.[2]
Another important reason for standardisation is the increasing globalisation. To be able to communicate and understand each other people and organisations for instance have to have common basis in order to know what their messages and exchanged information mean and how they can be trusted, managed and maintained. Good recordkeeping is an essential part of this.

This paper provides a short overview of the latest developments on standardisation in recordkeeping and discusses some of the underlying principles and key aspects of the current recordkeeping standards.

Work in ISO TC46/SC11

The objective of TC46/SC11 is to develop standards for records management. Its membership, however, is not limited to records managers or archivists, and consists of representatives from business companies, government organisations, professional associations, software vendors, and libraries. So far the committee has published three standards:

- ISO 15489:2001 records management
- ISO 23081-1:2005, metadata for records, principles[3]
- ISO 22310:2005, Guidelines for stating records management requirements in standards.

The subcommittee is rather active and has currently seven working groups, including

- 01: Metadata for records
- 02: RM relationships (suspended for the time being), which has published recently the ISO 22310 standard, providing Guidelines for stating records management requirements

[1] Members can only be standardisation bodies, but these are represented by stakeholders and people from interested organisations.
[2] See www.interpares.org, www.monash.edu.au/rcrg, www.naa.gov.au/recordkeeping/control/rkms/summary.htm.
[3] In 2004 published as a technical specification.

in standards. Since records management is often an inextricable part of doing business, this should provide guidance to other standardisation committees with respect to records management aspects in their standards.

- 03: Access rules. The objective of the group is to provide guidance on issues in relation with access to records, such as retrieval, interpretation, rights management, privacy, security.
- 04: Self-assessment/ compliance guide (suspended until the new version of the standard is published)
- 05: Review ISO 15489
- 06: Work process analysis. This working group has been established recently and will transform the Australian standard for work process analysis (AS 5090:2003) into an international standard.
- 07: Digital records preservation. This group will specify requirements for long term preservation. One of the objectives is to establish requirements for file formats to make them sustainable over time, instead of developing archival standards for each format, such as PDF. The work will be done in collaboration with another technical committee within ISO, ISOTC171/SC2, responsible for document management applications and imaging.

The latter two working groups will start in the Fall of 2005. The committee is not working in isolation, as may be clear from the above. It has relationships to many other related committees and organisations within and outside ISO. Within ISO there are many other committees and working groups (joint or otherwise) that do work very much related to the domain of records management.[4] TC171, responsible for standards in document management applications and imaging, is already mentioned. One of those standards that has been published in October 2005 is PDF/A.[5] Other committees are TC 10 that has delivered a standard for metadata related to technical documentation for instance, or TC 20 that has published the Open Archival Information Standard (OAIS).[6]
Especially in the field of metadata and interoperability collaboration is necessary. There is for example a liaison with a working group under the ISO/IEC joint technical committee (JTC1).[7] This JTC is active in areas such as EDI, metadata, database languages, and SQL/multimedia and has published standards on for instance metadata registries and framework for metamodel interoperability. Such standards will enable and support e-commerce and electronic service delivery.
Since TC46/SC11 has also relationships with organisations in the field of records and archives management, such as the International Council of Archives (ICA), the InterPares project and the International Records Management Trust (IRMT). Through these connections TC46/SC11 is gaining authority and influence and raising the status and the level of awareness with respect to records and archival management.

---

[4] Information about technical committees and related subcommittees can be found at https://www.iso.org/iso/en/stdsdevelopment/tc/tclist/TechnicalCommitteeList.TechnicalCommitteeList.
[5] ISO 19005-1:2005 PDF/A for PDF 1.4. The work is done in TC171/SC2/WG5 a joint committee with TC42, TC46/SC11 and TC130. Currently the PDF/A standard (ISO 19005-2) for PDF 1.6 is under way.
[6] The metadata for technical documentation standard is the ISO/IEC 82045 series; the OAIS standard is ISO 14721:2003.
[7] ISO/IEC JTC1 is responsible for standards on information technology. The committee meant is subcommittee (SC) 32 on data management and interchange, working group 2 which deals with metadata. See http://www.metadata-standards.org. They published for instance the ISO/IEC 11179 series, Information Technology - Metadata registries, and ISO/IEC 19763 series, Information Technology – Framework for metamodel interoperability.

The following paragraphs will focus be on the metadata standards and the review of ISO 15489 records management standard.

A standard for records management metadata

One of the main areas in records management is that of metadata for records. In 2004 a first technical report has been published, ISO 234081-1, that provides a framework for creating, managing and using these metadata and discusses the underlying principles. As such it is an extension of the ISO 15489 standard and explains what is necessary to ensure among other things the authenticity and integrity of records through time by using metadata. The definition says:

> "Metadata management is an inextricable part of records management, serving a variety of functions and purposes. In a records management context, metadata are defined as data describing the context, content and structure of records and their management through time (ISO 15489-1:2001, 3.12). As such, metadata are structured or semi-structured information that enables the creation, registration, classification, access, preservation and disposition of records through time and within and across domains."[8]

The standard explains furthermore what the benefits are for both business and records management processes to create and maintain metadata. Though different business contexts may require different approaches to metadata some basic principles can be identified as underlying the need for records management metadata. One of them is that metadata explaining the business context of records need to be captured together with the record itself, so it can be fixed in time and space (i.e. the domain in which the records are created and/.or used). Similarly, all records management processes performed on records after capture need to be documented through metadata. That information will support the authenticity, integrity, reliability and usability of the records over time. An important principle is also that every time a record is used again metadata about that new usage or business context need to be captured. It means that the record is never static or fixed, but keeps accruing metadata over time as it will be used or re-purposed over and over again as long as it exists.

In the following diagram the main entities are shown about which metadata should be captured.

[8] ISO 23081-1:2005, Clause 4.

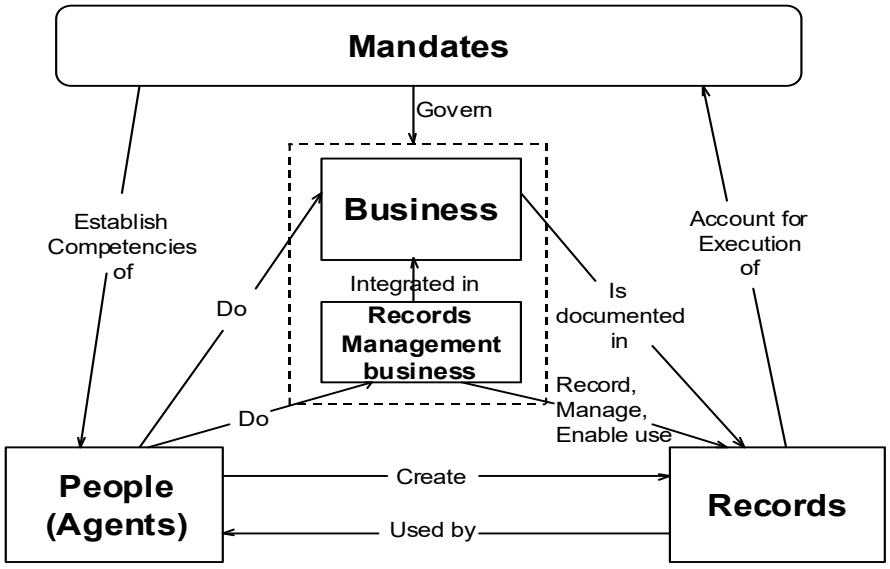

Figure 1 – Main entities and their relationships[9]

Basically, it reflects that people do business based on a certain mandate and in doing that create and manage records that document business transactions and can provide evidence of them. Records management is seen here both as integrated in doing business and as a business in itself with its own mandate and people doing it or responsible for it, e.g. managers, staff, or records managers. The standard explains for each of the entities what should be documented at point of capture of the records and after, providing thus a comprehensive overview of all the different aspects that need to be documented.

Apart from what should be documented the standard identifies the roles and responsibilities with respect to metadata. As may be expected, these are closely related to the responsibilities of records management.

Another area being discussed is the relationship with other metadata domains. Very important because there a lot of overlap with most other domains and vice versa. For example, metadata used in business processes should be captured to document why a record has been created and/or used, and that requires interoperability. Similarly, technical metadata necessary to reproduce digital objects and to enable their long term preservation is identified especially in the scientific and library communities, but is also relevant for records. Finally, the Dublin Core metadata set supports specifically the purpose of resource discovery on the web. This set overlaps entirely with metadata used for description of the records and subsequently with the records management objective to facilitate retrieval of records. Each of these sets of metadata elements has a different background purpose that has to be understood in order to see how they may or may not connect or used in other contexts.

[9] Derived from a report of a project of Monash University, called 'Recordkeeping Metadata Standards for managing and Accessing Information Resources in networked Environments over Time for Government, Commerce, Social and Cultural Purposes'. See http://www.sims.monash.edu.au/research/rcrg/research/spirt/reports.html.

It all touches on the issue of interoperability. In order to do business or to communicate it requires basically that information can be easily exchanged and subsequently used. That particular information needs to be equally understood by both parties involved. To achieve that aim it requires agreement about the structure and semantics of the exchanged information, such as about the metadata elements. For example, will the way an agent or organisation or also a business activity are described by one party be understandable to the other? This includes not only understanding the metadata schema and its (sub)elements, but also the encoding schemes that may specify the values of the elements, such as controlled vocabularies and thesauri.
It may also regard structural aspects of the metadata elements, such as the way a date is structured, e.g. DD/MM/YY or MM/DD/YY or even YYYY/MM/DD? In short, it is all about the meaning of things. That will differ between individuals, organisations, communities, and disciplines, because they will live and work in different business or social contexts. Each will have its own concepts and discourse. It will require mutual understanding and translation to bridge those differences.
Very much related to this is also the necessity to avoid duplication in work, that is recreating metadata every time again when documents or records are received by another organisation. An interesting research project currently undertaken in this area is the Australian Clever Recordkeeping Metadata project.[10] Its objective is to explore issues of interoperability and see how metadata created in one domain can be easily (re-)used in another, often for different purposes (re-purposing). These issues raise beyond the level of organisations or even communities and may require shared services, such as a metadata schema registry currently developed in the InterPares project.[11] It will require automated tools that can support the extraction of metadata and the translation of metadata elements between different metadata schemas from different environments (cross-walks), but also a thorough understanding of the metadata schemas. The latter may be difficult, since metadata schemas are not always very well documented.

The second part of ISO 23081 is still under construction and will zoom in on issues related to implementation. It will provide a further explanation why metadata are important in a business context and identifies policy issues and related responsibilities. The main part of this document will be dedicated to constructing a metadata framework and a schema. It addresses the underlying conceptual issues and subsequently how to design, construct and implement a metadata schema. In constructing a metadata schema organisations may use an existing similar metadata schema or parts of different schema and customise them to their own needs. This requires, however, a thorough understanding of why these metadata schemas are made. Certainly if an organisation constructs its schema based on parts of different other schemas. If not done carefully, the parts may not fit together well and thus may fail to serve the (new) purposes of the organisation involved.
The final section in this second part addresses issues such as appraisal and retention of metadata, capture, storage and preservation, presentation and finally the management of the actual metadata over time. It discusses questions whether metadata should be stored together with the record or separately. Presentation deals with human readable and machine readable expressions of metadata schemas. An example of a machine readable presentation is an XML-schema. Finally, this technical report will include a checklist of issues to be dealt with in relation to implementation of proper metadata including their management.

---

[10] See http://www.monash.edu.au/research/rcrg/research/crm/people.html. This project brings together researchers from Monash University and UCLA (USA), National Archives of Australia, State Records Office of NSW, and the Australian Society of Archivists (ASA). Also for more detail: Joanne Evans, Sue McKemmish and Karuna Bhoday, *'Create once, use many time: the Clever Use of Recordkeeping Metadata for multiple archival purposes'*, published at: http://www.wien2004.ica.org/imagesUpload/pres_174_MCKEMMISH_Z-McK%2001E.pdf.
[11] See www.interpares.org/ip2/ip2_description.cfm and Evans, McKemmish and Bhoday, 2004.

The third part of ISO 23081 will provide an evaluation instrument that will support organisations in assessing whether the metadata schema(s) they have adopted or constructed are compliant with the records management metadata principles laid down in part 1. This work is being done in close co-operation with the Description Group in the Inter Pares 2 project.[12] This group is developing a registry for recordkeeping metadata schemas. Examples of schemas included in the registry are the Recordkeeping Metadata Set of the Australian National Archives, the Minnesota records management metadata set, the ISAD/G, preservation metadata sets, and even the Dublin Core resource discovery set. The latter has been described because many people still think it may be adequate for records management purposes. The analysis shows it is not.

The project has developed also an evaluation instrument, that is built upon the principles in ISO 23081-1 and the authenticity requirements identified in the Inter Pares 1 project (1999-2001). Most of this work is done by people of the universities of Monash (Melbourne), Los Angeles (UCLA) and UBC (Vancouver). At the moment about 25 metadata sets and schemas are described of which 14 are analysed, while 26 are still waiting to be processed. The idea is that in the end the registry will be publicly available, but there are still organisational and financial issues to be resolved, especially the issue of sustainability. Who will maintain the registry, who will own the intellectual property rights, how will it be kept up-to-date? Part 3, however, will be an adapted version of the analysis instrument developed in Inter Pares and become an ISO technical report.

The whole series as such intend to offer a comprehensive set of documents, helpful for organisations in dealing with the issue of records management metadata.

The review of the records management standard ISO 15489

The other main work undertaken by ISO TC 46/SC11 is the review of the ISO 15489 records management standard itself. Within ISO every 3-5 years after publication of a standard a review has to take place, to see whether the standard is still valid and up-to-date.
The review process started in 2004 with a survey on the experiences with the standard so far. People were invited to comment on the standard and about 420 responses were received. In general, they were positive. It also showed that about half of the responding organisations is not (yet) using the standard. The comments will be used to scope and refine the standard. The main objective is not to come up with a completely new standard, but to make it more concise and clear in its phrasing and use of terms as well as to include some issues that are not well addressed in the current version. The latter regards for instance topics as access and metadata. The review will also take into consideration issues like what the audience should be, that it should be scaleable in the sense that it will be applicable for small as well as for large organisations. In principle the standard should be targeting in the first place senior management and not records management professionals. This will support the idea that records management should be seen as intrinsic to business activities. As a consequence records management has to be integrated in business processes. It is the intention that the renewed standard will be published at the end of 2006 or beginning of 2007.

The review will be done in conjunction with reviewing part 2 of ISO 15489, so the two documents will be better balanced. This part provides a further explanation of the standard and is targeting especially records professionals. Currently, some parts of this technical report belong in the actual standard and the other way around. An important issue is also the structure of the new

[12] See www.interpares.org/ip2/ip2_description.cfm.

version. Should it become a similar document as it is now or should it be completely restructured, e.g. into different technical reports. Also the content and the level of detail are not yet clear. The different working groups of TC46/SC11 reflect some of the topics currently in part 2. Decisions on these issues have a relationship with the activities and deliverables in other working groups of TC 46/SC11, since this can be seen as extensions and further explanation of the standard. The target date for the new part 2, regardless its structure, is at the moment 2008.

Some concluding remarks

The standardisation work in ISO TC46/SC11 reflects a certain international consensus within the domain on records management, but as it is developing through time it will require a periodical review to keep the standards up-to-date and useful. This is both required and supported by new insights and a growing understanding of the role and position of records management in new societal and business contexts that are rapidly changing under influence of information technology, increasing interconnectivity, and globalisation. Interoperability is the ‘magic’ word in this context, but we are only at the beginning of really understanding what the implications are and what solutions will be necessary. Standards that for instance explain the principles of metadata, how to construct schemas and how to document them will help and will provide a good basis, though.

The records management standards will provide organisations as well as the records management profession with a comprehensive framework that will enable them to manage their records for as long as required. In the field of records management they will help to raise the level of professionalism as well as the level of authority. The standards even reach out into the domain of archives, that is often still seen as a separate and distinct phenomenon, that of cultural heritage. However, the actual processes for managing the records are similar despite the fact that responsibilities may differ and the context may have expanded from an organisational to a societal perspective.

However, standards on their own do not resolve anything. They will help and support the actual business processes and the work of records and archives management, but in order to be effective they need to be included and embedded in business processes and in broader information and records management policies. This requires also more collaboration between disciplines and domains, including business, IT, and information management.