

Title: *International Recordkeeping/Records Management Metadata Initiatives: Report and Recommendations for DC Advisory Board*

Creator: John Roberts, Archives New Zealand

Creator: Andrew Wilson, National Archives of Australia

Date Issued: 2002-04-23

1. Background

The DC-GOV Working Group noted strong interest in records management metadata, or more broadly, resource management metadata in its working sessions at DC-2001 in Tokyo. There are a number of initiatives underway both within and across jurisdictions and communities that focus on this use of metadata, and may impact on the implementation of the DCMES. This discussion paper outlines the distinctions between this area of work and areas currently on the DCMI work plan, and considers possible roles for DCMI in responding to emerging needs in this area.

2. Metadata Initiatives

2.1 Discovery Metadata

The only significant global resource discovery initiative is the Dublin Core Metadata Initiative (DCMI). Discovery metadata, as the term is used in the DCMI, covers several distinct aspects of metadata functionality:

- location – enabling searchers to determine that a particular resource exists;
- evaluation – enabling some evaluation of the relevance and usability of the resource; and
- access – enabling the searcher to obtain the resource.

Discovery metadata does little to support the custodian of the resource in managing it over time. It is typically limited in its support of evaluation, dealing only peripherally with the reliability of the resource. However, discovery and access are essential components of recordkeeping and, ideally, discovery metadata elements will form a subset of a recordkeeping metadata element set.

The US Government Information Locator Service (GILS) was originally developed as a standard for resource discovery. GILS is briefly discussed below in section 2.3.

2.2. Administrative Metadata

Administrative metadata is metadata about metadata. It covers such matters as the author, date or currency of the metadata, and is a critical part of the effective management of a metadata-based system. Only indirectly, however, does such metadata support the management of resources themselves.

The influential *A-Core* draft (<http://metadata.net/admin/draft-iannella-admin-01.txt>) states: “Metadata about metadata - referred to as the A-Core - is useful to designate

information about the provenance, management or administration of other sets of descriptive metadata. The objective of A-Core is to provide simple verification of the integrity, ownership, and authorship of metadata retrieved from networked resources. The A-Core elements are utilised to associate the instruments (who, what) with the events (when) of the process of metadata management.”

2.3. Recordkeeping Metadata

In this paper the term ‘recordkeeping’ is used to refer to the process of “making and maintaining complete, accurate and reliable evidence of business transactions in the form of recorded information” (Australian Standard AS4390: Records Management, Part I-General). Recordkeeping includes records management and archives administration. The projects discussed below include both recordkeeping and records management metadata activities.

The recordkeeping community has been active in recent years in developing standards for recordkeeping metadata. Such standards are designed to support the management of the resource itself, as a record, over time. They thus support a wider range of functionality than just discovery. Typically, recordkeeping metadata initiatives have recognised the significant role played by DCMI, and considered how they can extend, rather than conflict with, the existing DCMES. We provide a definition of recordkeeping metadata sourced from the Archiving Metadata Forum (see: <http://www.archiefschool.nl/amf/>): “Structured or semi-structured information which enables the creation, management and use of records through time and within and across domains in which they are created. Recordkeeping metadata can be used to identify, authenticate and contextualise records; and the people, processes and systems that create, manage and use them.” Thus, recordkeeping metadata is crucial to the management of both current records (records management) and archives (archives administration).

Australia

The SPIRT Recordkeeping Metadata Research Project (<http://rcrg.dstc.edu.au/research/spirt/index.html>) has developed the Australian Recordkeeping Metadata Schema (RKMS), which provides:

- a standardised set of structured recordkeeping metadata elements;
- a framework for developing and specifying recordkeeping metadata standards;
- a framework for reading or mapping metadata sets in ways which can enable their semantic interoperability by establishing equivalences and correspondences that can provide the basis for semi-automated translation between metadata schemas.

The Recordkeeping Metadata Schema has been developed using conventions and protocols adopted by the wider metadata community, in particular the Dublin Core (DC: http://purl.oclc.org/metadata/dublin_core/) and Australian Government Locator Service (AGLS:

http://www.naa.gov.au/recordkeeping/gov_online/agls/user_manual/intro.html) metadata initiatives, to ensure compatibility, as far as practicable, between related resource management tools.

The National Archives of Australia (NAA) has developed the Commonwealth Recordkeeping Metadata Standard (RKMSCA: <http://www.naa.gov.au/recordkeeping/control/rkms/summary.htm>), which was issued

in May 1999. The NAA is now commencing a review of the standard with the aim of bringing it into closer alignment with the SPIRT schema.

The State Records Authority of NSW released in June 2001 its Recordkeeping Metadata Standard. See the NSW State Records website at: <http://www.records.nsw.gov.au/publicsector/erk/metadata/rkmetadata.htm>. This notes “Recordkeeping metadata is information that helps records to be identifiable, accessible and meaningful. It also supports the management of records for as long as they are needed. Capture and management of this information has traditionally been a routine part of records management. The majority of systems designed to control and manage records, such as automated records management products, should already capture this information. In these cases, the standard acts as a benchmark against which these systems can be measured and improved.” The NSW Recordkeeping Metadata Standard was influenced by both the NAA and SPIRT schemas and is closely modelled on the SPIRT schema.

Canada

The Canadian Government Information Management Forum has produced *Record Keeping Metadata Requirements for the Government of Canada* (http://www.imforum.gc.ca/new_docs/metadata3_e.pdf). This covers “the type of information ... required to capture the identity, authenticity, content, context, structure and management requirements of records created in the context of a business activity”. A number of the defined elements are mapped to DCMES elements.

New Zealand

Archives New Zealand is investigating the development of a recordkeeping metadata standard to support a range of initiatives around government recordkeeping practice. It is anticipated that work will be co-ordinated as far as possible with NZGLS, the New Zealand Government DC-based discovery metadata standard.

Sweden

A group of about 10 active members, led by Jonas Bosson is working on the development of a metadata standard for workflow processes (<http://www.wfmd.org/index.xml>). The group is aiming to develop a set of metadata elements for describing the processes a document is involved in and which will travel with documents as they move through electronic systems. This work is seen by the group as an adjunct to the records management metadata standard being developed by the ISO (See below). Document process metadata is, conceptually, a subset of recordkeeping/records management metadata, but also has links to preservation metadata, since the latter is interested in some of the aspects of documents that document process metadata would describe.

United Kingdom

The Public Record Office (PRO) *Functional Requirements for Electronic Records Management Systems* includes a metadata specification to support its functional requirements (Reference Document:

<http://www.pro.gov.uk/recordsmanagement/eros/invest/Reference.PDF>). The UK Government’s *e-Government Metadata Standard* (http://www.govtalk.gov.uk/interoperability/metadata_document.asp?docnum=524)

draws on both the DCMES and the PRO recommendations and is designed to meet both discovery and records management needs. It is a hybrid schema that contains only minimal records management metadata functionality.

United States

Although it is primarily a standard for resource discovery, the US Government Information Locator Service (GILS), now called the Global Information Locator Service, does include metadata elements that can be used for recording some aspects records management information about resources (see <http://www.gils.net/>). Originally developed for US Federal Government agencies in 1994, GILS now has a wider application and is enjoying strong support among US State jurisdictions. More information is at: <http://states.gils.net/>.

Between 1993 and 1996 the University of Pittsburgh's Department of Library and Information Science ran a research project to examine variables that affect the integration of recordkeeping requirements in electronic information systems. The project was called "Functional Requirements for Evidence in Recordkeeping" and was under the general leadership of Richard Cox. The principal outcome was the development of a "framework for business acceptable communications". This has been the major influence in the development of recordkeeping metadata policies and standards in Australia, Canada and the UK. A useful summary of the project and its outcomes by David Bearman, one of the project experts, can be seen at: <http://www.archimuse.com/papers/nhprc/BACartic.html>.

Current electronic records management initiatives in the United States are heavily influenced by the DoD5015.2 (<http://jtc.fhu.disa.mil/recmgt/>) standard for electronic records management systems. Jim Whitehead, chair of IETF's WebDAV working group, and Owen Ambur, co-chair of the XML working group, organised a meeting in November 2001 around the theme of how best to achieve interoperability of metadata among electronic records management (ERM) systems (<http://www.cse.ucsc.edu/~ejw/metadata/>). One of the outcomes was a proposal to develop an XML schema for records management metadata based on DoD5015.2. A draft schema is at <http://www.cse.ucsc.edu/~dgordon/ERM/ERMSchemaPaper.html>

International

There is ISO work currently in progress, under the leadership of Hans Hofman of the Netherlands, to develop an international recordkeeping metadata standard. The primary aim of this work is to identify the recordkeeping metadata needed to implement the international records management standard, ISO15489 (issued in late 2001). It is proposed that this standard also provide a framework for the development and implementation of specific metadata standards. Associated work of ISO TC46-SC11, which is responsible for the recordkeeping metadata work, is the investigation of the relationship of their work to other metadata initiatives. The work is little advanced as yet; a project plan has been drafted but not agreed on, or issued and the first meeting of the working group will take place at The Hague in late May 2002. No public documents are currently available (at 23 April 2002).

A related ISO initiative is the work of ISO Technical Committee 10. TC10 project IEC/CD 82045-2 is developing a set of management metadata for technical

documents. Part 2 of the project, a collection of metadata and reference models, is due by 30 June 2003. See

<http://www.iso.ch/iso/en/stdsdevelopment/techprog/workprog/TechnicalProgrammeTCDetailPage.TechnicalProgrammeTCDetail?COMMID=443>.

The University of British Columbia established a project, called InterPARES, in January 1999. This is described as “a major international research initiative”, involving participants from Canada, the US, Australia, Europe and Asia, whose goal is “to use the tools of archival science and diplomatics to develop the theoretical and methodological knowledge essential to the permanent preservation of electronically generated records. On the basis of this knowledge it will formulate model strategies, policies and standards capable of ensuring their preservation”. Since metadata is the obvious means by which the required information will be deployed, this project will, necessarily, address the issue of recordkeeping metadata standards, although this is not specifically listed on the research plan. See: <http://www.interpares.org/>.

A more recent project has been set up to develop an XML DTD for the Encoded Archival Context (EAC) specification. This project is being run from Yale University with the support of the Research Libraries Group, and an international committee of experts (<http://www.library.yale.edu/eac/>). The EAC project grew out of the EAD standard for describing archival documents, and is aimed at developing a prototype standard for representing descriptions of people, families, and corporate bodies. An EAC record contains elements for maintenance history, identity (of person, corporate body or family), relations between EAC records, resources relations, description of the functions or activities of the person or corporate body, systematic description of the entity and its environment, and biography or administrative history. Work to date is conveniently summarised on Robin Cover’s XML site at: <http://www.oasis-open.org/cover/eac.html>. EAC is at heart a metadata standard for describing agents, and as such has both discovery and recordkeeping applications.

2.4. Resource Management Metadata

Resource management metadata has a specific meaning in the Library community and relates to metadata needed for administration and management of electronic resources (such as e-journals) held by libraries. It is also a term that is used by the data management community to refer to the metadata needed to manage/control the data in large databases.

The tenor of the long discussions about resource management at both the DC-Gov and DC-Admin breakout sessions in Tokyo suggested that what was actually being referred to in a general way was what we have characterised in this paper as recordkeeping/records management metadata. However, not all resources need to be managed as rigorously as records, for which complex metadata standards are needed to ensure their maintenance and accessibility over time and space. A standard for resource management metadata could be developed which included discovery, records management, and preservation metadata but which would be less complex than a full recordkeeping metadata standard.

2.5 Other Related Metadata Initiatives

Preservation Metadata

There have been a number of projects in the last few years aimed at developing preservation metadata standards. Many of these projects have been modelled on NASA's Open Archives Information System (OAIS) which is probably the best known initiative (<http://www.ccsds.org/RP9905/RP9905.htm>). The OAIS defines the information required for preservation as: content information; representation information; preservation description information; and packaging information. A description of preservation metadata from the Preserving Access to Digital Information (PADI) page at the National Library of Australia website (<http://www.nla.gov.au/padi/topics/32.html>) states: "preservation metadata may be used to store technical information that supports preservation decisions and action, to document preservation action taken such as migration or emulation, to record the effects of preservation strategies, to ensure the authenticity of digital resources over time, and to note information about collection management and the management of rights". There are obvious links between recordkeeping metadata and preservation metadata, since preservation metadata is essentially a detailed expansion of one aspect of recordkeeping metadata.

Some preservation metadata projects specifically focussed on the metadata requirements for preservation of "born digital" resources:

The CEDARS project is run from the University of Leeds (UK). Its broad objective is to explore digital preservation issues. Those issues cover the acquiring of digital objects, their long-term retention, description, and eventual access. URL: <http://www.leeds.ac.uk/cedars/index.html>

NEDLIB is a collaborative project of European national libraries. It aims to construct the basic infrastructure upon which a networked European deposit library can be built. One of NEDLIB's projects was the development of metadata for long-term preservation. The report on this work is at: <http://www.kb.nl/coop/nedlib/results/D4.2/D4.2.htm>

The National Library of Australia's PADI project uses a preservation metadata set developed for the NLA's PANDORA project which aims to create a digital archive of significant Australian on-line publications. The draft PANDORA preservation metadata standard can be seen at: <http://www.kb.nl/coop/nedlib/results/D4.2/D4.2.htm>.

In the United States, the Online Computer Library Centre (OCLC) and the Research Libraries Group (RLG) have set up a joint working group on preservation metadata. This group "was formed in response to the need for consensus and convergence in the development, use and implementation of preservation metadata". The group's aims are to develop a comprehensive preservation metadata framework applicable to a broad range of digital preservation activities, and to examine issues surrounding the practical use and implementation of metadata to support digital preservation processes. More information is available from: <http://www.oclc.org/research/pmwg/>.

3. ROLE OF DCMI

In practice, metadata projects frequently need to consider metadata needs beyond discovery. At minimum, consideration needs to be given to how different types of metadata work interrelate – inconsistent expectations emerging from different communities jeopardises all work. The goal of interoperability does not relate solely to discovery across diverse metadata systems, but also of metadata compatibility between different metadata standards. This suggests a need for coordination and communication between development work within both the DC (resource discovery) and the wider recordkeeping/records management communities. At the very least it is essential for DCMI to have observer status with the more significant global initiatives.

Already the role of DCMI in records management has been contentious, with the Usage Board rejecting some aspects of a proposed DC-GOV application profile on the basis that they do not support resource discovery, but are records management oriented. The question of whether DCMI should become actively involved in developing a metadata set for resource management is a complex one, and relates to issues of the continuing role of DCMI within the global metadata community and the need for compatibility between metadata standards developed for different purposes (but all of which are fundamentally about managing aspects of information).

A number of possible mechanisms exist for accommodating work on the relationship of DC and resource management metadata within the DCMI.

Working Group?

DCMI Working groups are organized around specific problem domains. Working groups are formed and dissolved as dictated by the work at hand and the availability of expertise to accomplish such work. Working Groups require commitment from participants and the delivery of specific outcomes. They are focussed on achieving their aims in furthering the application and deployment of DCMES.

Interest Group?

Special interest groups provide a less formal mechanism than the Working Group model for discussion among members of the DC community on matters of common interest. Interest groups do not have specific deliverables or outcomes but function mainly as forums for discussion. Interest groups serve the need to bring interested parties together for discussion of common problems and can become working groups given sufficient impetus from participants and the desire to realise practical outcomes of use to participants and the wider DCMI community.

Extension of the role of an existing group?

Discussions on resource management metadata emerged at the working meetings of both the DC-GOV and DC-ADMIN working groups at DC2001 in Tokyo. This raises the possibility that any work in this area within the context of DCMI could be assigned to an existing working group. Although there are relationships to the work of both groups, this approach is not considered desirable.

While the DC-GOV working group explicitly discussed this matter, the implications are not limited to the government sector. Public sector recordkeeping is generally more regulated than in the private sector, providing a particular incentive for interest

from government practitioners in recordkeeping metadata. Other factors may focus attention in this area outside the government sector. In particular, recent cases of unauthorised destruction of records (the collapse of Enron in the US; a legal case against British-American Tobacco in Australia) have focussed the attention of the private sector on records management issues.

As noted above, we believe that resource management metadata is quite distinct from administrative metadata, being more complex and with broader application, and is better considered separately.

4. Recommendations

Effective development of compatible metadata schemas to support diverse needs of implementers requires effective communication between the various communities. DCFI as the first global metadata initiative serves as a model for metadata development for other communities. DCFI cannot afford to remain aloof from global metadata initiatives outside the arena of resource discovery.

We recommend:

- I. **Communication:** The DCFI Executive take steps as soon as possible to enhance communication between DCFI and global recordkeeping metadata initiatives. In particular, the Executive should seek involvement in the ISO recordkeeping metadata standard project, as an official observer if not an active participant. We believe that it might also be useful to seek some level of involvement with the Whitehead/Ambur project.
- II. **Information:** A member of the Advisory Board be tasked with following, and keeping up to date with, international and global recordkeeping/records management metadata initiatives, and reporting back regularly to the Advisory Board/Executive.
- III. **Participation:** DCFI establish a Working Group on Records Management Metadata to facilitate discussion and investigation within the DC community of:
 - the relationship of recordkeeping/records management metadata activities to resource discovery activities;
 - the integration of DCFI with recordkeeping metadata schema;
 - the possible involvement of DCFI in the development of a core set of resource management metadata elements.

John Roberts
Archives New Zealand

Andrew Wilson
National Archives of Australia