Internationaler Archivkongress 2004

23.-29. August - Wien - Österreich

Archive, Gedächtnis und Wissen

Smart Metadata and the Archives of the Future

Sue McKemmish

Joanne Evans

Anne Gilliland-Swetland

Nadav Rouche

Richard Marciano

Hans Hofman

Create Once, Use Many Times: The Clever Use of Recordkeeping Metadata for Multiple Archival Purposes

Joanne Evans, Sue McKemmish and Karuna Bhoday

Metadata is a key component in the creation, management and preservation of electronic records, as well as their innovative use as archives, memory and knowledge. However metadata generation and deployment are currently resource intensive and application specific. Metadata creation is not usually fully automated. Metadata created in one application of potential relevance to other applications is not shared between applications. Although data modeling, mark up language and syntax initiatives are addressing the data representation requirements for metadata translation and exchange, this functionality has not as yet been utilized in the systems that support eGovernment and eBusiness processes, electronic recordkeeping and archival description. Moreover there has been little progress in relation to developing strategies and meta-tools for the translation of metadata attributes and values between schemas in these environments. The Monash Clever Recordkeeping Metadata (CRKM) project addresses the challenge of automating metadata creation and sharing metadata between business systems, current recordkeeping systems and archival systems. This paper explores the relevance of the CRKM project to future archival systems and the deployment of metadata for multiple archival purposes. It is presented as part of the Smart Metadata and the Archives of the Future session that aims to communicate the progress and findings of several inter-related collaborative research projects and standards initiatives. Other papers in the session report on the related work of the InterPARES 2 Description Research Team (Designing a Meta-Registry for the Registration, Analysis and Archival Extension of Pre-Existing Metadata), the San Diego Supercomputing Center's development of Persistent Archives Technology (Metadata Tools and Sustainable Archives Technologies), and the ISO Metadata for Records Standard (Smart Metadata Research and International Standards).

The Archives of the Future

It is possible to re-imagine archival systems of the future that:

- Manage the records of multiple groups and individuals beyond the boundaries of the personal or corporate archive
- Represent multidimensional contexts of creation, capture, organisation and pluralisation – juridical, organisational, functional, procedural, technological and recordkeeping
- Provide multiple views of parallel recordkeeping universes
- Continuously and cumulatively weave relationships between records and related people, organisational structures, functions and activities to assist in preserving their evidential value and enable multiple access paths to records and their meanings
- Keep records relating to all recordkeeping and archiving processes persistently linked to the records they form and transform.

Such archival systems would have great potential utility in relation to the preservation and accessibility of electronic records of continuing value, as well as to the management of current records. The locus of the archives system might exist as an interface to archival records held by an archival institution, but it might also link to all records, publicly available or not, of continuing value or not (of continuing value), still maintained in the recordkeeping systems of individual agencies. In this sense, the collective archives could be preserved and made accessible in virtual space. Custodial arrangements and issues of where the record is physically located would cease to be of prime importance. (McKemmish et. al. 2005, Ch. 7).

Archivists at the beginning of the new millennium are challenged to develop archival systems for the globalised societies of the 21st century, systems that can operate beyond the level of the individual or corporate archive, and of collective archives as we now know them, to describe multiple recordkeeping realities, encompass the world views of all the parties to the transactions the records document, and provide meaningful access paths to all stakeholders. Chris Hurley has recently coined the term "parallel provenance" to refer to archival descriptive systems that could describe parallel recordkeeping universes:

> Recognising that the documentation created within the New Zealand national archives system largely reflects the cultural views of the Pakeha majority, but living in a society in which bi-culturalism is more than mere rhetoric, Hurley began to question how the views of the Maori could be accommodated in systems defined by Pakeha standards, and to seek a set of alternative, equally valid ways of viewing and documenting the records. He is currently exploring how the Australian

> series system, and related metadata schema developments, with their powerful relational characteristics might accommodate such differences, enabling alternative readings of the records and their contexts to be added by different communities of stakeholders. It is possible within these frameworks to represent records from different perspectives, from the point of view of the creator, other parties to the transaction, and other stakeholders, in and through time, from individual, community, corporate and societal perspectives. (McKemmish et. al. 2005, Ch. 7*)*

Eric Ketelaar has explored the implications of this approach to archival practice with reference to the concept of communities of records as developed by Jeannette Bastian, and associated ideas about shared ownership and joint heritage (Ketelaar 2004; Bastian 2003). Bastian defines a community of records as:

> the aggregate of records in all forms generated by multiple layers of actions and interactions between and among the people and institutions within a community (Bastian, 2003, p. 5).

According to this view, "the records of a community become the products of a multi-tiered process of creation that begins with the individual creator but can be fully realized only within the expanse of an entire community of records" (p. 3). Thus:

> all layers of society are participants in the making of records, and the entire community becomes the larger provenance of the records (p. 83).

Drawing out the implications of these conceptual approaches, Eric Ketelaar points to the matrix of mutual rights and obligations of all the parties involved and how they would extend to all aspects of recordkeeping and archiving – ownership, custodianship, appraisal, description, access and so on (Ketelaar 2004).

Essential to the development of archival systems of the future of the kind envisaged, systems that could also negotiate and manage such matrices of mutual rights and obligations, are emerging metadata management frameworks and schemas that specify the types of standardised information or metadata that integrated archiving and recordkeeping processes operating within broad archival frameworks would need to capture in order to fulfil these multiple purposes. Within these frameworks, recordkeeping metadata is defined as:

> structured or semi-structured information that enables the creation, management and use of records in and through time, and within and across the domains in which they are created and used. Recordkeeping metadata is used to identify, authenticate, and contextualise records, as well as the people, processes and systems involved in their creation, management and use (Wallace 2001, p. 255).

Metadata schemas provide semantic and structural definitions of metadata, including the names of metadata elements, how they are structured, and their meanings. Archival descriptive standards and control system specifications can be envisaged as traditional forms of recordkeeping metadata schema.

Within a records continuum frame of reference, standardisation of metadata and descriptive practices occurs across all recordkeeping environments, including business systems, recordkeeping systems, and systems for the long-term preservation of records of historical and cultural value (archives). Thus an emphasis on the clever use of metadata, including the re-use and inheritance of metadata from different business applications and environments for further utilisation in broader cultural and accountability domains, is an emerging and distinguishing feature of the approach of the Australian records continuum community of practice.

The Challenge

In current recordkeeping practice, many of the types of metadata created and used by records management systems are also created and used in a variety of other business applications, such as desktop document authoring, web content management, human resource management, and work flow systems. But records management systems as currently implemented do not draw on these other systems as sources of metadata; rather they re-create it – often in manual and resource intensive ways. A parallel situation exists in relation to resource discovery metadata. For example, in Australia AGLS resource discovery metadata, as specified in the national standard (National Archives of Australia

2004) is most often created retrospectively at the time a document is made available on a web site, rather than being automatically supplied by the software in which the document was originally created, although this application would have also created almost identical metadata in order to manage and retrieve the document. Archival organizations and programs also describe records of long-term value from scratch when they are transferred to their control, although David Bearman alerted us to the futility of such an approach and the need to develop strategies to enable descriptive systems to inherit metadata from current recordkeeping systems over fifteen years ago (Bearman 1989). Archival description, which assures the integrity and usability of records through time, is thus hugely resource intensive as it seeks to retrospectively describe records. Efforts to provide access to archival records through gateways that operate above the level of individual archival institutions and collections also rely on similarly resource-intensive methods of retrospectively describing the records with reference to metadata or archival descriptive standards designed to support such access.

If archival systems of the future are to provide for emerging concepts of parallel provenance and communities of records, and manage shared ownership, joint heritage and multiple access paths, the capture and implementation of metadata which can describe multiple contexts of creation, management and use in and through time will be essential. A key challenge is to develop systems, processes and tools that will enable metadata to be created once, then used and re-used many times for multiple purposes in different contexts in and through time.

Interoperability

The Clever Recordkeeping Metadata Project (CRKM) brings together researchers and practitioners from Monash University and UCLA, the National Archives of Australia, the State Records Office of NSW, and the Australian Society of Archivists' Descriptive Standards Committee to explore issues associated with metadata interoperability. The aim is to demonstrate how we can begin to move away from the current, resource intensive processes of manual metadata attribution and stand-alone systems, towards an integrated suite of business systems and processes supporting recordkeeping and archiving functions, environments in which metadata can be created once and used many times. It is this vision of integrated system environments and clever metadata which underpins much of the development in recordkeeping metadata standards to date.

Digital environments offer us the opportunity to move away from our "minimal descriptive systems" (Reed 2003, p. 19) – but we need to take the lead in articulating our requirements and demonstrating the "business" case for automating the capture and re-use of metadata required for recordkeeping and archiving processes, and reinventing archival description as a process of managing, augmenting, and re-purposing the rich mines of metadata in our environments.

There are many definitions of interoperability, but one of particular usefulness to us is that presented in discussion of the <indecs> metadata framework which provides a reference model relating to intellectual property rights management:

> Interoperability means enabling information that originates in one context to be used in another in ways that are as highly automated as possible (Rust et. al. 2000, p. 6).

With reference to this definition, we want to see metadata created in business systems which is relevant to recordkeeping being made available to records management applications, as well as metadata of relevance to business processes which is created in records management applications being re-used in business systems. We want to see metadata in organisational recordkeeping systems being inherited by archival control systems and metadata which originates in archival control systems being made available to recordkeeping and other business systems. We want to explore how metadata in archival control systems can be re-located into parallel recordkeeping universes or other information spaces. This involves exploring how metadata can cross technical, spatial and temporal boundaries in automated ways.

Towards Integrated Systems

Recordkeeping metadata standards developers look to a future in which their standards will be implemented in integrated systems environments that enable the clever use and re-purposing of metadata. In practice there have been significant implementation problems, as current systems environments do not as yet support the integrated processes for sharing metadata and re-using it for

multiple purposes anticipated by the standards developers. Moreover the development of meta-tools, such as metadata schema registries and mapping tools, which support the automatic creation of metadata and the translation of metadata attributes and values has not kept pace with the theoretical advances and standards initiatives. The metadata registries in the resource discovery sector, for example, have so far had limited functionality focusing on identifying and describing metadata schema and standards.
However, archivists are not alone in envisaging integrated system environments and some of the current practical limitations are being overcome by technological developments. The lack of support in current systems for metadata sharing and re-use is a legacy of closed, proprietary systems where any metadata or data exchanged with other applications would be hard-coded into the application. This barrier is being brought down by the trend to open systems and component based architectures[1]. For applications to be viable in such environments, their data and services must be capable of being accessed, invoked and manipulated by other system components. Standardized data representations across these components are also essential. This has led to the development of a number of technologies to facilitate this integration, e.g. encoding languages like XML for the representation of structured data, lightweight communication protocols like SOAP, and other technical standards supporting web services[2], as well as a proliferation of metadata standards initiatives supporting data interchange.

There is also growing interest in metadata schema registries as tools to support metadata re-use. A number of research and development projects have been undertaken or are underway to explore the architecture and functionality of schema registries to support metadata interoperability (OLIN; DCMI). The common vision is that such registries provide metadata about existing metadata element sets to software and/or human agents. With such metadata the agent can then determine the suitability and/or how to use the element set for its purposes. There is also interest in these registries managing mappings between metadata element sets and thus providing services to support metadata translation between applications.

Although these developments are of relevance to recordkeeping and archiving, a more comprehensive and complex view of the translation requirements associated with recordkeeping and archiving processes is emerging compared with that taken by other communities interested in metadata interchange. This relates to our need to examine how to make possible translations to and from business systems, recordkeeping systems and archival systems, translations across levels of aggregation, translations through time, and translations across contextual boundaries.

Layers of Interoperability Model

In 2001-2002 researchers involved in pioneering metadata registry activities in the information management area came together as a Working Group on Metadata Registries, sponsored by the DELOS Network of Excellence on Digital Libraries. Their aim was to consolidate their experiences and "articulate a shared set of principles underlying the construction of metadata registries" (Baker et. al. 2002). In the resultant white paper they present a simple model of the layers of interoperability in which issues associated with metadata re-use and re-purposing may be explored – see Figure 1.

Layer 3	**(a) Attribute Space** (e.g. LOM, Dublin Core, MES, indecs)	**(b) Value Space** (e.g. ontologies, classifications, controlled vocabularies, taxonomies)
Layer 2		

[1] 'Component technology is a blend of object-oriented and Internet technologies. In a component-based architecture, the components of a system have generic interfaces through which they advertise their functionalities, enabling the dynamic loading of the components' (Interoperability Clearing House, http://www.ichnet.org/glossary.htm)
[2] The W3C (World Wide Web Consortium) definition of web services is 'A Web service is a software system identified by a URI [RFC 2396], whose public interfaces and bindings are defined and described using XML. Its definition can be discovered by other software systems. These systems may then interact with the Web service in a manner prescribed by its definition, using XML based messages conveyed by Internet protocols.' http://www.w3.org/WAI/GL/Glossary/printable.html#W

	Representation (e.g. XML, RDF, DAML-OIL)
Layer 1	**Transport and Exchange** (e.g. HTTP Get, OAI Protocol for Metadata Harvesting)

Figure 1 Layers of Interoperability

From: Thomas Baker et al., *Principles of Metadata Registries*, 2002

Layer 3 is the abstract layer and is divided into an attribute space and a value space. Metadata consists of attributes, i.e. the characteristics or properties to be described such as Title, Date and Subject, and values, i.e. the values assigned to those characteristics such as the specific title, date and subject matter of a particular information resource. The attribute space encompasses the definition and description of the attributes and may be formally presented as a metadata standard. The definition of this attribute space may include data that identifies attributes, defines their purpose, describes usage rules and conditions and expresses relationships amongst them. It may also define the value space, i.e. the domains from which metadata values for an attribute may be sourced. This value space thus incorporates classification systems, controlled vocabularies and other instruments from which metadata values may be constructed. An example of this abstract layer in the resource discovery field is the definition of the Dublin Core Metadata Terms (available at http://dublincore.org/documents/dcmi-terms/). It specifies the elements, element refinements, encoding schemes, and vocabulary terms of the metadata schema. An example in the recordkeeping field is the National Archives of Australia's *Recordkeeping Metadata Standard for Commonwealth Agencies* (NAA 1999.1).

Underpinning this abstract layer is a conceptual data model. A conceptual data model identifies the "things" being described and their relationship to one another. In so doing it provides the perspective or the context of metadata, hence determining the characteristics of the attribute and the value space accordingly. For example, the Dublin Core Metadata Set is based on a bibliographic model where an information resource is viewed as an object, whose content has been created by authors, which is distributed by publishers, and over which certain rights are held. Hence these are the qualities of the object the metadata set seeks to capture. Whereas the NAA *Recordkeeping Metadata Standard* is concerned with information resources that function as records, the agents that create them, and the recordkeeping processes that manage them and assure their reliability and authenticity.

Layer 2 is the representation layer where "the attributes and values of Layer 3 are represented or instantiated using particular syntactic bindings in encoding languages such as XML or XML/RDF, which are processable by machines" (Baker et. al. 2002, p. 6). The relationship between layer 3 and layer 2 is therefore one to many, i.e. there may be many different representations of the abstract space. For example an XML Schema of the Dublin Core Metadata Element Set and an RDF Schema encoding for the Dublin Core Metadata Element Set with the definition of the title element expanded is illustrated below (see Figure 2 and 3 respectively).

The final layer incorporates the protocols for metadata transport and exchange that allow the representations from layer 2 to be moved between systems. This is the technical layer and as noted earlier there are many technological developments in this area making the integration of systems easier. The OAI Protocol for Metadata Harvesting is one such development enabling data providers to expose their metadata to service providers who can access it. For example the National Library of Australia uses this protocol in the provision of Picture Australia, "a federated discovery service based on aggregated metadata" about images from a number of Australian institutions (Boston 2003).

In the context of this model, metadata registries are defined as:

> applications that use metadata languages (Layer 3) in a form processable by machines (Layer 2) in order to make those languages available for use by both humans and machines. To be processable in automated ways, in other words, the conceptual structures must be bound to machine-processable formats (Baker et. al. p.7).

Metadata registries manage and manipulate representations of the attribute and value space, interacting with other systems using appropriate transport and exchange protocols. As a corollary to the above definition, if metadata registries are to support the automated translation of metadata between schemas, then the translations must also be in machine-processable forms.

```
- <xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns="http://purl.org/dc/elements/1.1/"
    targetNamespace="http://purl.org/dc/elements/1.1/" elementFormDefault="qualified" attributeFormDefault="unqualified">
  + <xs:annotation>
    <xs:import namespace="http://www.w3.org/XML/1998/namespace"
      schemaLocation="http://www.w3.org/2001/03/xml.xsd" />
    <xs:element name="title" type="elementType" />
    <xs:element name="creator" type="elementType" />
    <xs:element name="subject" type="elementType" />
    <xs:element name="description" type="elementType" />
    <xs:element name="publisher" type="elementType" />
    <xs:element name="contributor" type="elementType" />
    <xs:element name="date" type="elementType" />
    <xs:element name="type" type="elementType" />
    <xs:element name="format" type="elementType" />
    <xs:element name="identifier" type="elementType" />
    <xs:element name="source" type="elementType" />
    <xs:element name="language" type="elementType" />
    <xs:element name="relation" type="elementType" />
    <xs:element name="coverage" type="elementType" />
    <xs:element name="rights" type="elementType" />
  + <xs:group name="elementsGroup">
  - <xs:complexType name="elementType">
    - <xs:simpleContent>
      - <xs:extension base="xs:string">
          <xs:attribute ref="xml:lang" use="optional" />
        </xs:extension>
      </xs:simpleContent>
    </xs:complexType>
  </xs:schema>
```

Figure 2 XML Schema representation of the Dublin Core Metadata Element Set 1.1

Source: http://dublincore.org/schemas/xmls/simpledc20021212.xsd

```
  <?xml version="1.0" encoding="UTF-8" ?>
  <!DOCTYPE rdf:RDF (View Source for full doctype...)>
- <rdf:RDF xmlns:dcterms="http://purl.org/dc/terms/" xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  + <rdf:Description rdf:about="http://purl.org/dc/elements/1.1/">
  - <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/title">
      <rdfs:label xml:lang="en-US">Title</rdfs:label>
      <rdfs:comment xml:lang="en-US">A name given to the resource.</rdfs:comment>
      <dc:description xml:lang="en-US">Typically, a Title will be a name by which the resource is formally known.</dc:description>
      <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
      <dcterms:issued>1999-07-02</dcterms:issued>
      <dcterms:modified>2002-10-04</dcterms:modified>
      <dc:type rdf:resource="http://dublincore.org/usage/documents/principles/#element" />
      <dcterms:hasVersion rdf:resource="http://dublincore.org/usage/terms/history/#title-004" />
    </rdf:Property>
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/creator">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/subject">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/description">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/publisher">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/contributor">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/date">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/type">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/format">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/identifier">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/source">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/language">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/relation">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/coverage">
  + <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/rights">
  </rdf:RDF>
```

Figure 3 RDF Schema representation of the Dublin Core Metadata Element Set 1.1

Source: http://dublincore.org/2003/03/24/dces#

Recordkeeping Metadata Initiatives

The current status of metadata initiatives in the recordkeeping domain can be described with reference to this model. We can see that there has been much activity in the abstract layer with the development of a number of recordkeeping metadata standards and the specification of metadata needs in functional requirements for records management systems. In Australia, these initiatives have been part of a suite of standards, best practice models and guidelines developed to address the challenge of managing

electronic records and other information objects. The main objectives have been to support reliable, trustworthy and accountable business processes, to provide better access to information resources online, particularly in the areas of eGovernment and eBusiness, and to support the quality, and long term preservation and accessibility of the archival heritage (McKemmish 2000, Cunningham 2001). The suite includes international and national records management standards (Standards Australia 1996 and 2002; International Standards Organisation 2001 and 2004), and standards and guidelines issued in a number of national and state government jurisdictions, including resource discovery and recordkeeping metadata standards developed by two of the industry partners in the CRKM Project, the National Archives of Australia and the State Records Authority of NSW (NAA 1999.1 and 1999.2; State Records Authority NSW 2000; State Records of South Australia 2003; MoReq - Cornwell Management Consultants 2001; Public Record Office UK 2002).

Over a longer period of time, there have been a number of initiatives nationally and internationally to develop archival descriptive standards and models for descriptive practice. Examples include the work of the International Council on Archives Descriptive Standards Committee (http://www.ica.org/eng/mb.com/cds/descriptivestandards.html), and independent initiatives such as EAD, Encoded Archival Description (http://www.loc.gov/ead) and EAC, Encoded Archival Context (http://www.library.yale.edu/eac). These initiatives have also mainly addressed the abstract interoperability layer. The project's third industry partner, the ASA Committee on Descriptive Standards, has been a contributor to the development of standardised descriptive practices at both national and international level, particularly through the development and promotion of standards.

In Australia, the continuing evolution of the series system and the development of the Australian Recordkeeping Metadata Schema (McKemmish et.al. 1999 and 2000) as a framework standard for the development of sector specific recordkeeping metadata schemas and archival description standards "reach towards ways of representing records and their contexts as richly and extensively as possible":

> The Australian Recordkeeping Metadata Schema extends the Australian series system concepts of context, drawing on records continuum thinking relating to a record's complex and dynamic social, functional, organisational, procedural, and documentary contexts of creation, management, and use through spacetime, and informed by the insights of Chris Hurley and Terry Cook. The contextualisation provided in the Schema enables the linking of records to ever broadening layers of contextual knowledge in order to carry their meanings through time (McKemmish et. al. 2005, Ch. 7).

As indicated above, such initiatives relate to the abstract attributes and values layer of interoperability. In both the Commonwealth and NSW public sectors, concerns about the quality of current recordkeeping and the lack of compliance with standards point to the fact that so far the lower layers of interoperability identified in the model have not been addressed. For example, although Commonwealth government agencies are required to comply with recordkeeping metadata standards, and government web sites must use Australian Government Locator Service (AGLS) metadata to provide better access to government information online and facilitate eGovernment, audit findings confirm that in practice there are major problems with implementation of the standards (Australian National Audit Office 2002 and 2003). In relation to archival descriptive practice, the ASA Descriptive Standards Committee sees the development of processes and tools that support the clever use of metadata in archival description, in particular its inheritance from business applications and environments and its re-use for archival and cultural heritage purposes, as essential to the successful implementation of such standards.

Recordkeeping Metadata Representations

So while there has been much activity in the abstract attributes and values layer, there has been less in the representation layer. Developers of recordkeeping metadata standards or functional specifications for records management systems have tended to see representation issues as implementation issues. But we can see from the interoperability model, and its definition of metadata registries, that an essential requirement for automated re-use and re-purposing of metadata is the expression of the concepts in the attribute and value space in a machine processable form. So if recordkeeping metadata standard developers want automated metadata capture and re-use then they need to explore encodings that transcend particular implementations and foster the required interoperability. As part of this work, issues in the translation between the abstract and representation layers may also need investigation. The white paper of the Working Group on Metadata Registries notes that the constraints of an encoding

language may mean adaptation or even distortion of the underpinning conceptual models (Baker et. al. 2002, p. 7). Thus any such impacts will need to be assessed and addressed.

These representation issues are not for standard developers alone. It is the intent of the CRKM Project to foster dialogue between standards developers, implementers and vendors through the prototyping of an integrated system environment with tools to support metadata interchange. This environment can be used by all parties to gain shared understandings and insight into the issues that arise at this interface. For implementers and vendors, the progress made on the protocols for metadata transport and exchange means that they can now turn their attention from questions of "how to exchange metadata"to "what metadata to exchange".

Re-use and Re-purposing

The layers of interoperability model can also be utilised to locate issues associated with metadata capture and re-use that the CRKM Project is seeking to investigate.

In the attribute space, we need to explore how to describe and manage data about metadata schemas in order to allow for automated translations between metadata sets. For this we can draw on the work of a related project, InterPARES 2, outlined below, which is investigating the functionality of such a registry. As our focus is on metadata related to recordkeeping, we need to examine how to make possible translations between business systems and recordkeeping systems, translations across levels of aggregation, translations through time, and translations across contextual boundaries – a more comprehensive and complex view of translation requirements than that taken by other communities interested in metadata interchange at this stage. In the information management area the focus has been on transforming metadata across similar resource description systems, while discussions of metadata interchange in business systems relate to "in time" rather than 'through time' translations. The Semantic Web community is developing technologies to facilitate automated data sharing and re-use across boundaries, but their vision, at this stage, has been criticized as lacking understandings of cultural dimensions and unable to deal with changing meanings in time and place (Veltman 2004) We will therefore need to look at whether these initiatives can be extended to allow for the spatial and temporal translations of attribute spaces required to support recordkeeping.

Similarly we will need to examine how the value space can undergo multi-dimensional transformations. To move metadata values through time and space we need to investigate when and how to make the underpinnings and contexts of "knowledge organisation systems"[3] used to construct them visible and explicit in order to preserve meanings. We may also need to investigate how mappings between value spaces are made and defined.

As well as these issues in the abstract layer, we need to investigate and specify our requirements for methods and tools to support automated translation in the representation layer. Of particular relevance to this aspect of the CRKM Project are research initiatives being undertaken at the San Diego Supercomputer Center. These initiatives, as discussed below, have looked at tools and technologies for digital archiving and long term preservation, including metadata representation, translation and on-going management.

The CRKM: Conceptual Framework, Design and Methods

Conceptual Framework

The conceptual framework for the project is provided by records continuum theory and the records continuum model (Upward 1996 and 1997, McKemmish 2002), which support the development of complex, integrated systems and processes to manage records and archives in and through time, and across space. It has also provided the conceptual framework for the development of Australian records management and recordkeeping metadata standards, including the Australian RKMS, Recordkeeping Metadata Schema (McKemmish et al, 1999 and 2000), which provides a model for the development of sector specific recordkeeping metadata and archival descriptive standards, and is currently being redeveloped as an Australian national standard. The recently published ISO technical specification,

[3] The term *knowledge organization systems* encompasses the many tools used to classify, control and organize information, including authority lists, classification schemes, thesauri and ontologies. It was coined by the Networked Knowledge Organization Systems Working Group at its initial meeting in 1998.

ISO/TS 23081 -1:2004 Information and documentation - Records management processes - Metadata for records -- Part 1: Principles, also drew on this Schema, adopting its conceptual models to identify the types of metadata required to support the international records management standard, ISO 15489-1 (ISO 2001 and 2004). The ISO recordkeeping metadata standard, Australian RKMS and related sector specific schemas and their conceptual models are key components in the metadata management framework referenced by the CRKM Project.

Development of Proof of Concept Demonstrator

The first stage of the CRKM project involves the construction of a 'proof of concept' prototype to demonstrate an integrated system environment supporting automated metadata capture and re-use for a fictitious, though realistic, organisational setting. Within this environment, working models of the required meta-tools to facilitate metadata translations can be iteratively developed and evaluated. This stage will include the development and documentation of metadata creation, management, and multiple use scenarios as the basis for building the prototype.

The prototype will demonstrate how recordkeeping standards-compliant metadata can be created and captured through a mixture of automated and manual processes in particular application environments for subsequent re-use across applications for multiple purposes. It will be developed iteratively employing user-centred and rapid prototyping techniques. Existing office and workflow tools will be extended and integrated using scripting languages, to enable metadata to be re-used and value-added as the business transactions, records, information objects and resources that the metadata describes and manages move from one application to another in complex intranet and Internet environments, across domains and through time. This will allow investigation of issues associated with the translation of metadata across these boundaries as well as exploration of the design and functionality of supporting meta-tools. The aim is for the CRKM researchers to work closely with a programmer and a focus group of experts to extend existing software and metadata deployment functionality in small, user-centred iterations. Adopting this agile approach to systems development generates new ideas and re-prioritises old ones as the prototype evolves and insights develop. Unlike traditional systems development methods which attempt to develop a complete specification upfront, this approach is more responsive to evolving understandings of user needs and their interaction with technologies. (Martin 2003)

The Scenario

The National Archives of Australia will provide the test-bed for the first prototyping stage of the CRKM project, using a policy development - publishing - archiving workflow scenario, which is widely applicable to many organisations irrespective of jurisdiction or business activity conducted. Consequently, we expect the prototype based on this workflow will be an important demonstrator for engaging our audience and promoting interoperability.

The National Archives of Australia develops a range of resources from policies and strategies on recordkeeping in the Australian government to more detailed information, advice, standards, guidelines and manuals. In addition the Archives, publishes books and CD-ROMs about Australian history, genealogy and recordkeeping and guides to the collection. Using the selected scenario, the CRKM researchers will explore the creation, sourcing and capture of recordkeeping metadata in the process of developing such resources and creation of related records. They will then investigate how the recordkeeping metadata can be re-used for different purposes and in different environments:

1. to facilitate resource discovery as resources are published to the intranet and Internet, and
2. to facilitate archival intellectual control given that the 'master copies' of these resources and records relating to their development and publication form part of the national archives.

The Prototype

Working with a focus group from the National Archives and guided by the Australian work process analysis national standard, CRKM researchers will derive recordkeeping and metadata requirements and specifications that will inform the iterative prototyping work. The focus group will assist CRKM researchers by helping to establish that the workflow is correctly identified and documented, source relevant authoritative metadata, validate specifications and iterations of the prototype as the prototype evolves, and provide valuable evaluation on the business utility of metadata and on the feasibility of integrated systems supporting automated capture and re-use of metadata. The first iteration of the

prototype will examine automated capture and re-use on a very small scale to demonstrate that the concept is possible. For example, looking at the underlying metadata standards for the prototype, the project may demonstrate translation of a few elements across the standards with subsequent iterations building on this work.

The prototype's simulation of systems and records will be derived from the business application environment of the National Archives. There policies are developed using desktop authoring applications, such as MS Word and Outlook. Records are captured and managed in a TRIM records management system. Resources are published to the public website and internally to the intranet. 'Masters' of these resources and related records will also be transferred to form part of the national archives and intellectually controlled by *RecordSearch*[4]. The proof-of-concept demonstrator prototype will be based in this business application environment, but will using the most current version of TRIM, and will simulate an archival control system using *Tabularium*[5] rather than *RecordSearch*. It is envisaged that subsequent iterations of the prototype may involve adding the functionality of content management systems to the environment, and extending the scenario to involve the translation of metadata from the archival control system to a higher level gateway to archival resources, such as those developed using the *OHRM*[6] system of the Australian Science and Technology Heritage Centre.

Implementation Model

The second stage of the project will involve implementing the prototype in a real world test-bed site to provide a model for best practice. Two different contexts for implementing the prototype will be used – the context of a single organisation and the context of the performance of a single function or activity across multiple organizations. This stage will include specification of the functional requirements for the implementation of the prototype at the test-bed site, extending the user-centred, iterative approach to systems development taken in the prototyping stage.

Meta-registries and Meta-tools

Prototypes of tools for metadata translation and deployment will be developed or adapted to support both the prototyping and implementation modelling stages. A working model of a mini metadata registry will be built to document and describe the different metadata standards or schemas employed by the different business applications used in the scenario and test-bed site. Its functionality will be extended to enable translation of metadata between the different metadata schemas in use so that metadata can be re-used for different business purposes and in compliance with different standards.

In developing the registry, representations of metadata schemas in machine processable forms will be investigated, drawing on the outcomes of recent data modelling, representation and syntax initiatives, as well as related work undertaken in the InterPARES 2 project and at the San Diego Super Computer Center, as described below. Mappings between the attribute and value spaces of metadata schema and mappings between metadata schemas and metadata held in business systems will also be developed. Depending on the boundaries across which these translations will need to occur, these mappings may involve rules for the aggregation of data or for making contextual metadata explicit. Investigation of how these mappings can be represented for automated processing will follow, in the first instance in the form of machine processable crosswalks, and then through iterative study of the suitability of emerging semantic web technologies.

During development of the proof-of-concept demonstrator prototype, the project will examine and experiment with the HotMeta suite developed by the Distributed Systems Technology Centre (http://www.dstc.edu.au/Products/metaSuite/) as the means of supporting metadata interchange; and consequently, additional tools will be developed to support this functionality. The prototype will explore how recordkeeping metadata compliant with the Recordkeeping Metadata Standard for Commonwealth Agencies can be re-used as resource discovery metadata compliant with the AGLS

[4] *RecordSearch* is the archival control system of the National Archives of Australia, see http://www.naa.gov.au/the_collection/recordsearch.html.
[5] *Tabularium* is a free-ware collection management system for archives based on the Australian series system model, http://tabularium.records.nsw.gov.au/.
[6] OHRM is a context based resource discovery and access system that links creators, archival and heritage resources and published materials based on the Australian series system model, see http://www.austehc.unimelb.edu.au/ohrm/.

Metadata Standard and re-used as archival descriptive metadata compliant with the CRS (Commonwealth Record Series) Manual, the implementation manual for the National Archives' implementation of the series system. The project will develop the tools needed to support this translation. This will include manual mappings (in both directions) between the various metadata standards and later machine readable cross walks to enable automated translation. The project will also develop where necessary the XML DTDs and XML Schemas needed by the HotMeta suite to support the metadata standards.

Relationship to Other Research Projects

InterPARES 2

The CRKM Project has a close formal relationship with a major international research initiative, InterPARES 2, the International Project on the Preservation of Authentic Records in Electronic Systems, 2002-2006[7] , which is researching the reliability, authenticity, preservation and accessibility of electronic records of artistic, scientific and government activities:

> InterPARES 2 aims at developing a theoretical understanding of the records generated by interactive, dynamic, and experiential systems, their processes of creation, and their present and potential use in the artistic, scientific and government sectors. On the basis of this understanding, the project will formulate methodologies for ensuring that the records created using these complex systems can be trusted as to their content (that is, are reliable and accurate) and as evidence (that is, are authentic), … selecting those that have to be kept for legal, administrative, social or cultural reasons after they are no longer needed by the organizations that created them, preserving them in authentic form over the long term, and analyzing and evaluating advanced technologies for the implementation of these methodologies in a way that respects cultural diversity and pluralism (SSHRC Grant Application Document, InterPARES2, http://www.interpares.org/).

Outcomes from the work of the Description Team within InterPARES 2 will feed into the CRKM Project, particularly the development of a Metadata Schema Registry, which is particularly relevant to the translation of metadata attributes and values in the abstract layer of the interoperability model, and research relating to meta-tools. Similarly outcomes from the CRKM Project will also be used by Description Team researchers.

San Diego Supercomputer Center

A number of SDSC initiatives that are building on knowledge in the metadata and meta-tools research space are directly relevant to the CRKM Project. The San Diego Supercomputer Center's (SDSC) Data and Knowledge Systems (DAKS) program is of particular interest to the CRKM project. The DAKS program (http://daks.sdsc.edu/) includes two research labs and a number of research projects that are of mutual interest to the CRKM project. The Sustainable Archives and Library Technologies (SALT) Laboratory http://www.sdsc.edu/Press/03/012604_SALT.html) and the Data Grids Technology Group are investigating a number of interoperability issues. The SALT lab will be developing information technology strategies and conducting research in the areas of digital archiving and long-term preservation. The Data Grids Technology Group is pursuing research that supports interoperability across distributed resources. The Storage Resource Broker (SRB) is an example of the type of technology that is being developed that supports automated metadata management. The SRB, in conjunction with the Metadata Catalog (MCAT), provides everywhere access to data sets and resources based on their attributes rather than their names or physical locations (http://www.npaci.edu/DICE/SRB/).

Research into electronic records management sponsored by the National Archives and Records Administration (NARA) and SDSC on issues relating to preservation and metadata archiving has led to

[7] Funded by the Canadian Social Science and History Research Council, SSHRC, Can$4m), InterPARES is made up of national teams from Europe, the UK, Asia, Africa, Australia and North America. The CI in the CRKM Project (Professor Sue McKemmish) is a member of the InterPARES International Team and Director of the Australian National Team, while PI Associate Professor Anne Gilliland-Swetland is a member of the International Team, and Co-Director of the US National Team. McKemmish and Gilliland-Swetland are also Co-Directors of the International Team's Description Group responsible for coordinating InterPARES research relating to standards for recordkeeping metadata and archival description.

a number of other research projects aimed at developing infrastructure technologies and tools for long term preservation of digital objects. In particular, research into the feasibility of a 'persistent archive'; a term used by the SDSC researchers in some of these projects to describe the data bits comprising digital entities and also the context that defines the provenance, authenticity, and structure of the digital entities (Moore, 2003). Another archival project of particular interest, Methodologies for Preservation and Access of Software-dependent Electronic Records (also known as Archivist Workbench, http://www.sdsc.edu/NHPRC/) is looking at the feasibility of developing prototypes of useful tools for archivists to preserve and provide access to electronic records over the long-term. Building further on this research is a project; called the Persistent Archive Testbed that will enable participating archival institutions to test SDSC's data grid and persistent archives technologies on a variety of archival collections. The research undertaken in these projects is highly relevant to the interoperability issues being explored by the CRKM project, particularly those relating to the representation layer of the model.

Research and Standards Initiatives

As has been discussed throughout the paper, there is a close interrelationship between the development of recordkeeping metadata standards, including archival descriptive standards, which address the abstract attributes and values layer in the interoperability model (Layer 3), and the research being undertaken in the CKRM Project. The recently published ISO technical specification, *ISO/TS 23081 - 1:2004 Information and documentation - Records management processes - Metadata for records -- Part 1: Principles*, provides a benchmark for the development of best practice metadata standards in the records and archives field relating to recordkeeping metadata attributes and values, and a set of criteria against which existing and future schemas and standards can be assessed in terms of their recordkeeping and archiving functionality. The research being undertaken by the CRKM Project will feed into the further development of the ISO, including its possible extension to address representation layer interoperability issues relating to the complex metadata translation requirements of recordkeeping and archiving processes.

Conclusion

The records continuum provides a conceptual framework that enables simultaneous multiple views of recordkeeping "realities". The Australian series system and more recent initiatives like the Australian Recordkeeping Metadata Schema and the ISO standard for recordkeeping metadata set up frameworks for capturing layers of rich contextual metadata, and multiple contexts of creation, management and use in and through time to support archival purposes in the democratic societies of a globalised world of the kind envisaged by Terry Cook, Tom Nesmith, Verne Harris, Michael Piggott and Sue McKemmish (Cook 2000, Piggott and McKemmish 2002, Nesmith 1999, Harris 2001).

Most archival system implementations currently privilege the role and rights of the records creator, and represent the records creator's world view and context rather than the world views and contexts of other parties to the transaction, drawing on functional classification schemes developed at the level of the corporate or individual archive. However, within these emerging frameworks we can look to the further development of metadata and description strategies which will enable archivists in the 21st century to go beyond Scott's original vision of sequential multiple provenance to build archival systems of the future that can encompass Chris Hurley's "parallel provenance" and Jeannette Bastian's communities of records, and negotiate the complex matrices of mutual rights and obligations that Eric Ketelaar's vision of shared ownership and joint heritage invokes. Within these frameworks the translation of metadata between business, recordkeeping and archival systems, across levels of aggregation and contextual boundaries, and through time will be of critical importance. The research on methods and tools to support such translations being undertaken in the CRKM Project, InterPARES 2 and at the San Diego Supercomputer Center will contribute vital understandings and strategies to the clever use of recordkeeping metadata in forming and transforming the archives of the future.

ACKNOWLEDGEMENT

This research was funded by an Australian Research Council Linkage Grant (2003-5), the National Archives of Australia, the State Records Authority of NSW, the Descriptive Standards Committee of the Australian Society of Archivists, and Monash University. Chief Investigators of the Clever Recordkeeping Metadata Project are: Professor Sue McKemmish (Monash University); Associate Professor Anne Gilliland-Swetland (UCLA) and Adrian Cunningham (NAA). Other personnel include

Research Fellow Karuna Bhoday (Monash University and NAA), PhD researcher Joanne Evans (Monash University), agile programmer, Sergio Viademonte, research coordinator Carol Jackway, industry partner researchers Tony Leviston (SRA NSW), Barbara Reed (ASA) and Duncan Jamieson (NAA), and DSTC research scientist, Dr Andrew Wood. For more information about the Project, see http://www.sims.monash.edu.au/research/rcrg/research/crm/index.html.

REFERENCES

Australian National Audit Office (2002), *Recordkeeping (Audit Report No. 45, 2001-2002)*, 1 May 2002, http://www.anao.gov.au/WebSite.nsf/Publications/4A256AE90015F69BCA256BA5000C25D8

Australian National Audit Office (2003), *Recordkeeping in large Commonwealth agencies (Audit Report No. 7, 2003-2004)*, 24 September 2003, http://www.anao.gov.au/WebSite.nsf/WhatsNew/478B4A27724E193BCA256DA50074796C!OpenDocument

Baker, Thomas, Blanchi, Christophe, Brickley, Dan, Duval, Erik, Heery, Rachel, Johnston, Pete, Kalinichenko, Leonid, Neuroth, Heike, and Sugimoto, Shigeo, (2002) *Principles of Metadata Registries*, A White Paper of the DELOS Working Group on Registries, DELOS Network of Excellence on Digital Libraries, 2002, http://delos-noe.iei.pi.cnr.it/activities/standardizationforum/Registries.pdf

Bastian, Jeannette (2003), *Owning Memory: How a Caribbean Community Lost its Archives and Found Its History*, PhD Thesis, published in 2004 as part of the Contributions in Librarianship and Information Science Series

Bearman, D. et al (1999), "A Common Model to Support Interoperable Metadata: Progress Report on Reconciling Metadata Requirements from the Dublin Core and INDECS/DOI Communities", *D-Lib Magazine* 5:1 (January 1999), available at http://www.dlib.org/dlib/january99/bearman/01bearman.html

Bearman, D. (1989), *Archival Methods* (Pittsburgh: Archives & Museum Informatics, 1989)

Boston, Tony (2003), "National Library of Australia Initiatives using the Open Archives Initiative Protocol for Metadata Harvesting", A paper given by Tony Boston, Director, Digital Services at a National Archives of Australia *International Seminar on the Use of Standards in the Development of Online Access Systems for Archives,* Canberra, October 2003, available at http://www.nla.gov.au/nla/staffpaper/2003/boston1.html

Cook, T. (2000), "Beyond the Screen: The Records Continuum and Archival Cultural Heritage", keynote address at Australian Society of Archivists Conference Melbourne, August 2000 http://www.archivists.org.au/sem/conf2000/terrycook.pdf

Cornwell Management Consultants (2001), *Model Requirements for the Management of Electronic Records (MoReq)*, (INSAR Supplement VI, Office for Official Publications of the European Commission, Bruzelles-Luxembourg, March 2001), available at http://www.cornwell.co.uk/moreq.

Cunningham, A. (2001), "Six Degrees of Separation: Australian Metadata Initiatives and Their Relationships with International Standards", *Archival Science* 1:3 (2001) 271-283

DCMI, Dublin Core Metadata Initiative, Open Metadata Registry, http://rdf.dev.oclc.org/myrdf/services/EOpenRegistry

ISO (2001), International Standards Organisation, ISO 15489 *International Standard: Records Management*

ISO (2004), International Standards Organisation, ISO/TS 23081-1:2004 *Information and documentation - Records management processes - Metadata for records -- Part 1: Principles*

Harris, V. (2001), "Of Fragments, Fictions and Powers: Resisting Neat Theorising about 'The Record' " in *Convergence* (Australian Society of Archivists and the Records Management Association of Australia, 2001), 11-18

Ketelaar, E. (2004), "Communities of Records", Paper presented at the School of Information Management and Systems, Monash University Melbourne, Research Forum, 23 July 2004

Lagoze, C. and Hunter, J. (2001), "The ABC Ontology and Model", http://metadata.net/harmony

Martin, Robert (2003), *Agile Software Development: Principles, Patterns, and Practices*, (Pearson Education, New Jersey: 2003)

McKemmish, S., Piggott, M., Reed, B. and Upward, F. (2005), *Archives: Recordkeeping in Society* (Charles Sturt University Press, Australia: publication pending)

McKemmish, S. (2002), "Placing Records Continuum Theory and Practice", *Archival Science* 1:4 (2001) 333-359.

McKemmish, S. (2000), "Collaborative Research Models: A Review of Australian Initiatives", *The American Archivist* 63:2 (Fall 2000) 353-67

McKemmish, S., Acland, G., Cumming, K., Reed, B. and Ward, N. (2000), *Australian Recordkeeping Metadata Schema (RKMS), Version 1.0, 31 May 2000,* http://www.sims.monash.edu.au/research/rcrg/research/spirt/deliverables.html#arkms

McKemmish, S., Acland, G., Ward, N. and Reed, B. (1999), "Describing Records in Context in the Continuum: The Australian Recordkeeping Metadata Schema", *Archivaria* 48 (Fall 1999) 3-43

Moore, R. and Merzky, A. (2003), "Persistent Archives Concepts", *Global Grid Forum* (December 2003) available at http:www.npaci.edu/DICE/Pubs/Data-PAWG-PA.doc

National Archives of Australia (1999.1), *Recordkeeping Metadata Standard for Commonwealth Agencies Version 1.0* (Canberra: NAA, May 1999), available at http://www.naa.gov.au/recordkeeping/control/rkms/summary.html

National Archives of Australia (2004), *AGLS Metadata Element Set, Version 1.3* (Canberra: National Archives of Australia), available at http://www.naa.gov.au/recordkeeping/gov_online/agls/metadata_element_set.html

Nesmith, T. (1999), "Still Fuzzy, But More Accurate: Some Thoughts on the 'Ghosts' of Archival Theory", *Archivaria* 47 (Spring 1999), 136-5

OLIN, UK Office for Library and Information Networking, DESIRE Metadata Registry, http://desire.ukoln.ac.uk/registry/

Piggott, M. and McKemmish, S. (2002), "Recordkeeping, Reconciliation and Political Reality", in *Past Caring*, *Proceedings of the Australian Society of Archivists Conference, Sydney 2002*

Public Record Office (2002), *Functional requirements for Electronic Records Management Systems*, UK Public Record Office, available at http://www.pro.gov.uk/recordsmanagement/erecords/2002reqs/

Reed, Barbara, "Recordkeeping in Business Systems", *Informaa Quarterly* 19:3 (August 2003), 18-22.

RCRG Monash (1998), Archival metadata standards for managing and accessing information resources in networked environments over time for government, social and cultural purposes; Chief Investigators Sue McKemmish and Anne Pederson, Principal Investigator Steve Stuckey; funded by an Australian Research Council SPIRT Grant and industry partners State Records NSW, National Archives of Australia, Queensland State Archives, Australian Council of Archives and the RMAA, http://www.sims.monash.edu.au/research/rcrg/research/spirt/about.html

Rust, Godfrey and Bide, Mark, (2000) *The <indecs> Metadata Framework Principles Model and Data Dictionary* (June 2000), available at http://www.indecs.org/pdf/framework.pdf

Standards Australia (1996), AS 4390-1996 *Australian Standard: Records Management* (Homebush: Standards Australia)

Standards Australia (2002), AS 15489-2002, *Australian Standard: Records Management* (Homebush: Standards Australia)

State Records Authority of NSW (2000), *NSW Recordkeeping Metadata Standard (NRKMS), October 2000*, available via http://www.records.nsw.gov.au/publicsector/erk/metadata/rkmetadata.htm

State Records Authority of New South Wales, *Government Recordkeeping Manual*, includes Standard on Full and Accurate Records, Manual for Design and Implementation of Recordkeeping Systems (DIRKS), and policies on Electronic Recordkeeping and Electronic Messages, available at http://www.records.nsw.gov.au/ via NSW Public Sector link

State Records of South Australia (2003), *South Australian Recordkeeping Metadata Standard (SARKMS)*, http://www.archives.sa.gov.au/publications/SARKMS%20Metadata%20Scheme%20Final%20AMW%20290503.pdf

Upward, F. (1997), "Structuring the Records Continuum Part Two: Structuration Theory and Recordkeeping", *Archives and Manuscripts* 25:1 (May 1997), 10-35

Upward, F. (1996), "Structuring the Records Continuum Part One: Post-custodial Principles and Properties", *Archives and Manuscripts* 24:2 (Nov 1996), 268-85

Veltman, Kim H. (2004), "Towards a Semantic Web for Culture", *Journal of Digital Information* 4:4, (March 2004), http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Veltman/

Wallace, David, "Archiving Metadata Forum: Report from the Recordkeeping Metadata Working Meeting, June 2000", *Archival Science* 1:3 (2001), 253-269