

Metadata: Core Record or Core Business?

Barbara Reed

Barbara Reed is a Senior Lecturer in the Department of Librarianship, Archives and Records at Monash University. Prior to joining DLAR in 1995, she worked as a recordkeeping consultant in the private sector and government records environments for over 10 years. Her research interests focus on electronic recordkeeping and the strategic integration of recordkeeping into business processes. She has a BA (Hons) from the University of Sydney, a GradDipArchAdmin from New South Wales and an MA (Hons) from Melbourne.

This article raises critical questions about the way we should be managing the metadata associated with records and recordkeeping processes in order to maintain records in their context through time in complex and rapidly changing environments. It explores some current models for specifying record metadata, drawing on the outcomes of research projects and standards initiatives, and speculates about the usefulness of identifying a core set of record metadata, given the contingent nature of recordkeeping activity. In mapping the overlap between the metadata specified in the University of Pittsburgh and University of British Columbia projects, and the Australian Records Management Standard, the paper uncovers a possible 'core' set of record metadata. Analysis of that core reveals that it would essentially enable the description of the record as a passive object. This discovery leads to a questioning of the viability of passive approaches to records description and a proposal that the process of appraisal be reconceptualised. The paper goes on to outline the enormous challenges involved in managing meaning through time, focusing on issues associated with complex and inter-related cultural, functional and structural changes in organisations. It concludes by suggesting that

it is critical to adopt an active and dynamic approach to specifying, attributing and managing the metadata associated with records and recordkeeping processes, one that takes into account the contingent nature of recordkeeping activities. This would in part involve recordkeeping professionals in building on the outcomes of recent metadata-related initiatives at the same time as they move beyond the boundaries set by those initiatives.

This is a refereed article.

Introduction

Metadata is one of the current ‘buzz words’ in the discussion of how to manage electronic information. It is a powerful concept, but misunderstood and misapplied, like any concept, it becomes meaningless. In essence, metadata is simply information about information. It is a word that has been borrowed from the information technology field where it originated. Because the term is so general in application, we need to be careful to explain how we are using it. To use the term loosely and without qualification is dangerous, as it is very easy to confuse metadata which belongs to different layers of aggregation – the information entity itself, accumulations of related information entities, or systems designed to control the information entities.

It is easy to discuss records in terms of metadata. All of our traditional records and archives control systems are metadata management systems: that is, they are designed to manage information about records – information such as the registration details, the classification codes, the location details. Some of these metadata elements are *included* in the record itself and extracted to form part of the metadata management system: the form or document type, the date, the author etc. Others are *derived* from the record into the metadata management system during the performance of a recordkeeping process (such as précis, location codes, linkages or relationships to other records).

When faced with implementing systems in the electronic work environment which will act as records systems – that is, will provide evidence of business transactions – we need to re-visit many of the propositions which were taken

for granted or assumed by the ways in which we have managed paper records. We need to be able to make explicit the metadata elements required to manage documents as records.

In managing electronic information, the opportunity to capture appropriate metadata must be integrated into the process which creates and maintains the record in order to maximise the benefits deriving from the technologies. In any system, retrospective capture of information about events, transactions or use is not desirable, affordable or feasible. Electronic systems are less forgiving than paper systems: if we do not accurately define and capture the metadata needed to manage the record and make it accessible over time, that record in effect becomes lost in a data swamp – unconnected to the context of its creation, left without the essential information needed for its interpretation, and unlocatable in the mire. Failing to attribute as much metadata as possible from creation and management processes, and the contexts of creation and use, will severely limit the capacity of recordkeepers to provide value added information. Where inadequate metadata is attributed in the first place, recordkeepers are condemned to attempt the tedious and resource-intensive process of attributing this fundamental metadata retrospectively, rather than enabled to enhance, enrich or add value to access and retrieval processes, and the transfer of meaning over time.

The challenge before us is to specify:

- what metadata should be appended to an electronic document to make it a record;
- by what agent (human or machine);
- at what point in its existence;
- in response to which events or triggers; and
- under what information architecture models.

In order to do this, we must understand why we do this.

- Is it to create a representation of a record-object, a surrogate of the record?

- Is it to represent a set of data about the processes of recordkeeping?
- Is it to enable the management of the record over time?

Once we are clear on these issues, we also need to consider what we must do to ensure that metadata will retain its capacity to be interpreted over time.

Recordkeeping as a contingent activity

We have answers to many of the questions about record metadata, but as is always the case when dealing with records, the answers can often only be defined in relation to the particular environment in which the record needs to be kept. In other words, recordkeeping is essentially a contingent activity. The concept of a record as evidence of business activity is now well established. Defining exactly which transactions require the full evidential protection of recordkeeping, and the extent of the metadata which needs to be recorded to ensure this, are dependent upon the particular industry, the professional 'best practice' and the societal context in which the specific organisation works.

In some highly accountable organisations (and these exist both within the private and the public sphere), every transaction in key business functions must be protected as evidence of those transactions. Examples include the management of nuclear fuel in the power industry, ministerial correspondence, treaty negotiations, authorisation of foreign exchange transactions or share market trading. The degree of evidential protection required is not dependent on how long the record needs to be kept, rather it is related to the accountabilities associated with the particular function being carried out. The nature of a particular society or organisation, as well as the expectations of those with a legitimate interest in the transactions, also drives it. These expectations relate to being able to receive a true account of the commitments which affect their rights and interests, and which have been undertaken in their name.¹

This essential feature of recordkeeping – its contextual and contingent nature – makes it futile to establish hard and fast rules. This, in turn,

frustrates those who turn to the recordkeeping profession for rules which can be applied universally in implementing information systems which will act as responsible recordkeeping systems. Our answer, 'Well, it depends ...' is often taken as a sign that we do not know. On the contrary, it is the only response possible. Recordkeeping solutions do not come out of a box. They are, and must be, tailored to the particular requirements of individual organisations. Analysis and thought is needed to determine the recordkeeping responsibilities of organisations. Recordkeeping must be tailored to the requirements of specific business functions and activities, linked to related social and legal requirements, incorporated into particular business processes, and maintained through each change to those processes.

Proposal of a 'record core'

The detail of what metadata to record about documents, in order to ensure that they can act as evidence of transactions, is therefore not hard and fast. Some records will require the full weight of evidential protection, ensuring that whenever a query is raised there is full documentation available to defend the authenticity and reliability of the record. In other cases, the chances of being called to account are slim and perhaps organisations can run the risk of not having the same recordkeeping protection. This will impact upon costs of implementation – it may not be worth implementing full recordkeeping protection for all records.² The risks may be less than the costs potentially involved.³

There is a view that the metadata which constitutes the core of any record can be specified and standardised.⁴ This minimal level of specified metadata would then allow the record to be defended as an accurate representation of the business transaction. The core itself would not contain all the elements required to evidence the continuous processes of recordkeeping, ie to prove the record has been maintained and managed in accordance with recordkeeping best practice over time. Using such a minimal core, however, recordkeepers might be able to retrospectively reconstitute the circumstances of its maintenance and management and thereby defend it as evidence of business transactions. What is added beyond the record core would then be up to the particular application, business activity and organisation to define in terms of the risks likely to be associated with not

having a fully documented record of particular transactions.⁵ The viability of developing a standardised core set of record metadata, given the contingent nature of recordkeeping, still needs to be established.

The models

Various models have been put forward recently to specify what metadata is essential to ensure that a document will act as evidence of transactions. Comparison of these models reveals very different approaches to specifying, attributing and managing record metadata. Nevertheless, it is possible to map the models against each other to identify the extent to which the elements common to each might constitute a core set of record metadata.

The models are:

- the templates associated with the University of British Columbia's 'Preservation of the Integrity of Electronic Records' project
- the Business Acceptable Communications (BAC) model of David Bearman derived from the University of Pittsburgh's 'Functional Requirements for Evidence in Recordkeeping' project
- the Registration section of Part 4 of AS 4390 *Australian Standard on Records Management*.⁶

Each of these proposals comes from different perspectives. The first two models have both emerged from academic research projects, whereas the third was grounded in identifying and standardising 'best practice'.

The University of British Columbia (UBC) project derives its model from the application of the science of diplomatics to electronic records and contains language and concepts specific to that discipline. It is based on a traditional view of records and record systems of the custodial registry kind. While using a very specific and particular language, the template proposals also contain many elements which are familiar to many of us as we grapple with the emerging electronic environment. The templates focus attention on the record itself, and do not take account of the recordkeeping processes that manage it.

The Business Acceptable Communications model is a communications model, premised around the transmission process and how that transactionality fundamentally defines a record. It assumes an electronic context; specifically an object oriented environment – a particular information architecture that encapsulates metadata in layers around the record content. Each layer can be opened progressively according to permissions. The metadata swaddling the record is presented in a logical sequence which instructs the receiving or opening software about each of its attributes, conditions and characteristics in the order in which they need to be processed or understood. The object oriented approach looks to the concept of an intelligent document carrying with it all the metadata needed to contextualise and interpret the record over the whole of its physical existence. The language used to express the metadata specifications derives from the information technology world and is therefore more difficult for most recordkeepers to come to grips with.

The proposal in the Australian Standards document is modest and not on the same research level as the other two proposals.⁷ It specifies only two required metadata elements and suggests additional elements. It does not assume any particular information architecture. It clearly focuses on the record and derives from traditional records management practices.

In its expansion into a non-exclusive set of descriptive information it ventures into the electronic world. It is included here, because, despite its modesty, it contains all of the elements common to both of the research models and is expressed in Australian, a language more familiar to us than the language of either diplomatics or object oriented information analysis.

The models are very different in terms of the language used and the rationale behind their development. Comparison between the models is therefore very tricky and not regarded by their proposing project teams as particularly useful. Nonetheless, interpreters of the projects, and those grappling with how best to implement electronic records solutions, have consistently suggested a synthesis of the models in order to find a 'record core', from which particular extensions can be developed according to social and organisational requirements for recordkeeping in particular contexts.

A 'record core'

Table 1 is a preliminary analysis of the common elements in these three models undertaken to explore the possibility of identifying an initial listing of what might constitute a 'record core'.

As noted above, undertaking such an analysis reveals the clear differences in philosophy that are reflected in the models as proposed, and affect their ability to be mapped. It should also be noted that interpretation of the concepts of both research models, as presented here, might be contested.

It is easier to plot the UBC templates against the registration detail of AS 4390 than against the BAC model. Both UBC and AS 4390 take a view of item level registration which is essentially passive. The view is that of the record as passive object – something to be described. The BAC model implies a far more dynamic object, one that embodies, enables and encompasses recordkeeping processes.

In mapping the models, the elements of the UBC templates and the AS 4390 specification can be adequately 'fitted' together. Significant numbers of the BAC model's mandatory elements cannot. These additional elements are listed in Table 2. They are associated with recordkeeping process data. Their inclusion in the mandatory requirements of the BAC model, in both the terms and conditions layer and in the use history layer, implies a different view of the record-object than that taken by the other models. These process-related metadata specifications contain elements which enable the record-object to resolve permissions for access, use and disposition, and accept future metadata relating to the transaction history of who has accessed the record and when. By contrast, the record core derived from the overlapping parts of the three models, as represented in Table 1, enables only the description of the record as a passive object.

The degree of specificity of metadata that we need to include in the registration part of the record core also depends upon where the record is managed. Within the domain of an organisation, everyone accessing the data brings with them a common understanding that the record exists within the organisation. In traditional paper based records management systems we tend not to define this organisational understanding.⁸ In an environment

TABLE 1

Comparison of metadata from 3 models

| University of British Columbia templates relating to electronic records ¹⁸ | University of Pittsburgh BAC model specifications ¹⁹ | AS 4390, Australian Standard on Records Management, Part 4, Registration section ²⁰ |
|---|--|--|
| Medium* (assumption that a record must be recorded information fixed onto a media base is common to all although only specified in UBC) | (assumed) | (assumed) |
| Content* | Content Layer (V.A) | (assumed) |
| Transmission* (assumption that a record must be transmitted is common to all models although only specified in UBC) Mode ¹ ; security protections ¹ eg cryptographic seal; date stamping; corroboration ¹ | (This is the basic premise of the model – it is a communication model) File Encoding Metadata (III.B.1-5) | (assumed) |

TABLE 1 (cont)

| | | |
|--|---|--|
| <p>Form* comprising physical form and intellectual form</p> | <p>Record Declaration (I.A.1) File Rendering Metadata (III.C.1-5) Record Rendering Metadata (III.D.1-2) Content Structure Metadata (III.E.1-6) Business Transaction Type (IV.A.4)</p> | <p>viii) physical form xi) application software and version under which the record was created or in which it was captured xii) standard with which the records structure complies xiii) details of embedded document links including applications software and versions under which linked record was created xiv) templates required to interpret document structure</p> |
| <p>Persons*: entities the juridical system recognises as having the capacity to act, incorporating name and signature of author^{oo}, name of addressee^{oo}, names of receivers^{oo}</p> | <p>Transaction Context Metadata (IV.A: originator identification, recipient identification) Source Metadata (III.F)^Δ Data Source (III.F.1) Data Capture Instrument Type (III.F.3) Data Capture Instrument Settings (III.F.4) Responsibility Metadata (IV.B) Originating Organisation (IV.B.1)</p> | <p>v) author vi) sender vii) recipient</p> |
| <p>Date (includes both place and time)* Incorporating time and date of transmission^{oo}, time of receipt^{oo}, topical date^{oo}</p> | <p>Record Identification Metadata(I.A) Transaction Domain Identifier (I.A.2) Transaction Instance Identifier (incorporating date time and ID sequence) (I.A.3)</p> | <p>b) date and time of registration iii) date of creation iv) date and time of communication and receipt</p> |

TABLE 1 (cont)

| | | |
|--|---|--|
| Act* (connection with an action) Disposition [∞] | ? Action requested (IV.A.7) <i>not mandatory</i> | ? ix) links to related records documenting the same sequence of business activity |
| Title or subject [∞] | ? Information Discovery Content Metadata (1.B.1-3) <i>not mandatory</i> | i) document name or title ii) text description or abstract |
| Registry number [†] | <i>included in Transaction Instance Identifier 1.A.3)</i> | a) a unique identifier assigned by the system |
| Archival bond* Classification code [†] | File Identification Metadata III.A File Linking Rule/Standard (III.D.1) Set Relationships (III.E.8) Dynamic Relationships (III.E.9) Linked Prior Transaction (IV.A.6) | ix) links to related records documenting the same sequence of business activity x) business system from which the record was captured |

* Data elements from Template 5.

∞ Data elements from Template 6.

† Data elements from Template 7.

^f Data elements from Template 8 (note: template 8 addresses systems issues not incorporated into the individual record but used to protect the authenticity of a record – although not linked to the individual transaction, these might be linked to the BAC metadata specified in IV.C System Accountability Metadata).

Δ Data elements may or may not represent juridical persons. To 'map' to the UBC specifications, they would have to be juridical persons.

where the organisational boundaries are ever changing and unclear, such as in networked systems or cyberspace, we need to specify this basic domain identification because we can no longer guarantee a shared understanding of the domain of creation or use. A record core metadata set which lacks such specificity, detailing only requirements for a unique identifier, will not support interpretation of the record outside the creating domain. To enable that, we need more detailed specification of the domain itself, data which is redundant when you know where you are, but essential to understanding and interpreting records where the domain is not explicit.⁹

If we assume for the moment that we can arrive at an agreed mandatory core set of metadata, which will ensure that documents have the basic capacity to act as evidence of transactions, how would such a core be employed? The University of British Columbia project proposes templates as part of document creation. The AS 4390 proposes recording such details as a part of the recordkeeping process of registration, while the BAC model implies automatic capture of such data as records cross communication boundaries. The applicability of a record core, and where and how it could be implemented, needs further thought and exploration.

Beyond implementation issues, we also need to look at what might not be in the record core that has traditionally been part of the extended set of record metadata:

- What purpose did that additional metadata play?
- What would we lose if we adopted a record core?
- If a passive representation of a record is appropriate as the record core, what can we do with it?
- What additional metadata would have to be included to enable recordkeeping processes?
- Where and how should these processes add their metadata to the record core?
- How vital is that process metadata?
- What different information architectures or different recordkeeping system models would be needed for the different approaches?

- How will the validity of the information incorporated in the metadata of any record core or extended set be maintained?

Recordkeeping processes and metadata

It is not enough simply to create and capture a record. That record must be managed once it has been captured. In the course of that management, it attracts additional metadata elements. Indeed, as Chris Hurley argues:

Data captured by a records management system is not so much about the record as it is about the recordkeeping process. The entities 'described' by records management systems are not simply objects waiting to be depicted. They are actually components of a recordkeeping process through which the described object (the record) passes ... In the same way that a record is a representation of a business process, the description of a record is a representation of a recordkeeping process.¹⁰

Insofar as the set of metadata identified by mapping the three models constitutes a record core, we can see that it primarily is made up of the metadata that is associated with the recordkeeping process of registration. It seems acceptable to all engaged in thinking about electronic recordkeeping to assume that the registration of items as records can be undertaken by the records creator, or the records creator's system, at the time of creation. That is, there seems to be agreement that this process, once adequately specified with key metadata, can be delegated out to the point of creation – in effect ratifying the notion that every records creator is involved with the recordkeeping function.

While concentrating primarily on the registration process, the UBC templates and the AS 4390 specification also venture into classification as a prerequisite part of the metadata required.¹¹ But, if data associated with classification is brought into the set of data attributed at the time of creation, why not incorporate data needed to perform other recordkeeping processes?¹² Such recordkeeping processes include appraisal, access and use, storage (or representation) and migration. It is this logic that the BAC model embraces (see Table 2).¹³

TABLE 2

**MANDATORY ITEMS FROM BAC MODEL
NOT ABLE TO BE MAPPED**

(NB grouped by recordkeeping process,
not BAC model order of presentation)

| | |
|-----------------|---|
| Access | Access Rights Status (II.A.1) Access Conditions Resolver (II.B.1) Resolver Terms (II.B.2) |
| Use | Use Rights Status (II.A.2) Use Conditions Resolver (II.C.1) Use Terms (II.C.2.a-c) |
| Disposal | Removal Authority (II.D.1) Retention Policy Citation (II.D.2) Retention Period End Time (II.D.5) |

The BAC model provides the framework for specifying a record container which enables processes to be performed on the record-object. The framework allows the record to be embedded into its context of creation, in order to disembed it from any particular application. Once swaddled in its metadata, the record is able to exist in any environment as it carries with it the data needed to extract, exclude or allow access to the content and to make that content comprehensible. Things still need to happen to the encapsulated object: permissions need to be resolved, additional metadata attributed, migration triggered etc. These are recordkeeping processes, but within the different information architecture embodied in the model, they become facilitative actions. The systems or programs or applets acting as triggers do not have to carry meaning over time. The meaning is embedded with the encapsulated record. The record metadata is itself the entity which records and carries the meaning of those transactions over time. This allows us new ways of thinking about what recordkeeping systems and processes do.

A part of the process of appraisal is addressed implicitly by all models. But we need to extend our notions of what appraisal means in relation to electronic records. Decisions about what records to capture as evidence of business activities and at what point to capture them are presumed to be made prior to the metadata registration process. Associated with that initial appraisal decision to capture a record, however, comes the broader appraisal process of determining what metadata needs to be attributed so that the record will be managed over time. Appraisal drives the decision processes involved in determining what metadata to include with, or as a part of, the record. If we do not adequately specify the metadata associated with ensuring the carriage of the record through time, appraisal decisions are taken out of our hands. Appending that metadata, depicting the record, is the recordkeeping process of description.

To manage records for the period that they need to be retained, we need to incorporate decisions about disposal or how long the record should be maintained. Such decisions can be incorporated at the point of metadata attribution by reference to an external set of warrants, legislation, regulations and statements of best practice in relation to maintaining records (analogous to traditional disposal authorities). Disposal triggers

need to be attached to records in a systematic way, using rules embedded in authority tables.¹⁴ At the time of destruction, is authorisation still required? Do we destroy the content of the record and its metadata or the content only, leaving the metadata to act as evidence of the fact that a record existed?

Specifying how long records need to be retained also involves decisions about how they will be accessed. Who is permitted to access the records of particular business transactions, in what form, and how long should the restrictions imposed apply? As a corollary of the access permissions, in many situations we also need to maintain a record of who has used the record and in what further transactions the record has played a part.

Appraisal decisions also need to embrace the notion of migration. At some point in any electronic recordkeeping system, there will be a need to migrate records whose software or hardware dependency is becoming unsustainable. In our record core we have documented these dependencies. But, we also need to incorporate some type of monitoring of dependencies, with triggers to alert the recordkeeping system that migration is now required, and a mechanism to invoke the migration process. These processes should also provide additional metadata on the date of the migration and a system check on the reliability of the migrated record.

In passing, it is important to note that none of these recordkeeping processes, whether conceptualised as separate processes or encompassed within the appraisal process, can yet be appropriately satisfied by software currently available. The issues of access and event history or usage history are the most advanced of these in both electronic document packages and in electronic records management packages.

Recordkeeping assumptions in attributing record metadata

The aim in proposing a record core is to define those minimum metadata elements that are necessary to ensure evidence of transactions. In developing and implementing such a specification, it is clear that some recordkeeping assumptions will already be givens. For example, prior to the attribution of metadata elements, we still have to work out what documents need to be

captured as records, at what point and by what agent. Linking the documents to the business activity assumes a prior definition of business functions and activities. There is also a presumption (although only implicitly built into the record core) that the person who is doing the transaction is competent or has the authority to transact the business. None of these things can be assured by merely capturing metadata specific to one record. These are organisational issues, which must be resolved separately and precede specific instances of metadata attribution.

The models all assume transmission or communication as a basic element of the record, which provides us with some clues about events which might trigger or be devised to 'trap' records as they pass through defined communication boundaries. These communication boundaries will, however, need to be defined in relation to the particular business activities being undertaken, and implemented prior to the creation of records in order to ensure the circumstances which will result in the capture of a record.

Managing meaning in metadata

If we can agree on a viable record core – and this paper is suggesting that such a core would need to be more extensive than the set of elements common to the three models analysed here – we still need to face the challenge of the metadata itself being time bound. This is a necessary corollary of it being a record. In order to capture the record appropriately it must be bound to the context of its creation and not change. However, organisations and the language we use to describe their activities are not static things. They must change. Ensuring that the time and context bound record continues to be able to be interpreted in a changed and changing environment requires the establishment and maintenance of complex sets of relationships over time.

Some of the elements which will need to be monitored in order to ensure the appropriate interpretation of the metadata are those describing:

- the individual and their competence;
- the business function and activity; and

- the language used to describe the activities documented in the record.

This is not a comprehensive list, as it may also be necessary to establish cumulative relationships between templates and other forms. However, the three more obvious examples illustrate the point.

We need to monitor individuals and their competence. It is essential to know who carried out an action and whether they were authorised by the organisation to act in the particular matter. Detailing only the name of an individual is not going to provide any protection of authority. Names are not unique: there may be many J. Bloggses in an organisation – one may be the cleaner and one may be the chief executive officer. The authority to act (competence), as vested in the different positions held by people called J. Bloggs, is determined by the position held and the accountabilities incorporated into the organisational role being undertaken. An individual's name is therefore by itself an incomplete identifier. A position name also needs to be included and this needs to be linked to statements of accountabilities, authorisations etc. We also need to have this information on a cumulative basis. If J. Bloggs, the cleaner, is later promoted to become the Information Technology Manager, we will need to be able to link the records created at the time that J. Bloggs was the cleaner to that role. It is the particular role and organisational responsibilities which define the meaning and authority of the action embedded in the record. Simply updating organisational records to indicate that J. Bloggs is now the Information Technology Manager will invalidate the interpretation of any earlier records created by J. Bloggs in his previous roles.

Business functions and activities will also alter with time. Again, because the record must be time bound, it must capture with it the description of the business at the time of the transaction it records. But over time, as organisations reorganise and reengineer their business functions and activities (particularly their activities), the boundaries of the activities alter, as do the names they are known by. This is now a familiar characteristic of every organisation. Linkages need to be made between records created in one business activity and those created in the successor business activity. The records of particular business activities are the inheritance of the successor activities and, if still in existence, they are presumably needed for

current management purposes. To alter the way the record referred to the business activity of which it formed a part would be to invalidate the record. We need to capture changes to organisational structure through time in ways which will clearly identify links between existing records and present structures. (In the paper world, we did this by physically splitting and relocating parts of file series at times of administrative change – Australian government records managers have been particularly adept at managing this process.)

In the same way we need to accept that concepts change. Over long periods of time the examples of these changes are clearer. We only need to think of the words used to refer to Aboriginal and Torres Strait Islander peoples in early twentieth century government and church systems to realise that what might have appeared to be acceptable terminology to describe individuals or characteristics then is now understood to be as deeply offensive. Yet to alter those offensive words in documents and file titles will change the record and take it out of the context of its creation. Shifts like this occur more subtly all the time. Personnel becomes Human Resource Management; EDP becomes computing; vogue words and abbreviations come in and out of fashion. Records managers have developed expertise at managing the shifts in language in organisations through tools like thesauri and controlled vocabularies. Rather than rewrite the previous description, sophisticated linkages to ensure connections between terminology over time are required.¹⁵

In short, managing metadata and the meanings inherent in metadata are issues which also require attention. To perform recordkeeping tasks over time, we are not able to dispense with recordkeeping systems. They are needed to trigger actions associated with recordkeeping processes of appraisal, migration and use. In addition, they are needed to maintain the cumulative knowledge bases which are essential to ensure that metadata retains its ability to render meaningful records in context, while at the same time maintaining their connections to the changing work environment. As Chris Hurley has argued in relation to allocating descriptive terminology, an external point of reference is always necessary:

Metadata essential to an understanding of a record (x created the record) must be comprehensible (who is x?). Knowledge of context

could conceivably be encoded but understanding cannot. Understanding depends upon contextual knowledge, which is also historical and thus must necessarily exist outside the record.¹⁶

An exploration of the externality required to interpret metadata and meaning over time is critical to the reconceptualisation of the purposes of the archival information system.

Conclusion

The demand for certainty, uniformity and hard and fast rules in recordkeeping ignores the reality that recordkeeping is a contingent activity, depending upon its context and the risks tolerable in each organisation. We are not going to reach uniformity – but we might achieve agreement about a minimum set of metadata attributes. However, we need to be alert to the danger that, by accepting too minimalist a representation of records, we might be accepting a limited vision of how recordkeeping systems work, one that may not be viable into the future, a passive rather than an active view of records. In particular, if we attempt to develop a core set of record metadata from models that see records as objects to be acted upon, rather than active participants in business processes and technologies, our efforts may be futile. But, if we take a more active, dynamic approach, the contingent nature of recordkeeping must be explicitly addressed. Before we pursue proposals for deriving a record core from existing models and specifications, we need to understand the recordkeeping assumptions inherent in the models and the issues they raise. The models of recordkeeping we adopt will also drive how metadata attribution is implemented.

In pursuing our understanding of these matters we must be free to explore what the underlying conceptualisations of recordkeeping are. At the same time as protecting and guarding those elements of our recordkeeping tradition which ensure evidence of transactions, we must discard the practices which seek to emulate the way we did things in a paper based world. In looking towards electronic futures, we need to make sure that we can see beyond a first generation electronic notion of records – the equivalent of

the precursor to the car being the horseless carriage.¹⁷

How we conceptualise our recordkeeping practice defines how we specify metadata. If records are passive objects to be described retrospectively, we will take a different view of the what metadata is appropriate than if we see record metadata as accumulating from a set of recordkeeping processes which can recur a number of times over the record's existence. The preliminary analysis of the three models reported here suggests that we could derive a record core of common elements, but how useful would it be? Are passive models the right models? What could we do with the record core derived from a synthesis of the current metadata statements? The answers to these questions will reverberate through the implementation options adopted.

Endnotes

- ¹ This is a point of agreement between two of the major research projects into electronic records – the University of Pittsburgh project into Functional Requirements for Recordkeeping, which addresses the issue under the concept of 'literary warrant' (<http://www.sis.pitt.edu/~nhprc/>) and the University of British Columbia project, The Preservation of the Integrity of Electronic Records, which uses the concept of juridical environment (<http://www.slais.ubc.ca/users/duranti/>).

- ² The recent Electronic Records Research Working Meeting held in Pittsburgh reflected participants' concerns about the lack of hard facts about costs associated with implementation options, a lack matched by the dearth of implementation projects actually being undertaken in the workplace. See D. Bearman and J. Trant, 'Electronic Records Research Working Meeting, May 28-30, 1997. A Report from the Archives Community' in *D-Lib Magazine* July/August 1997, (<http://www.dlib.org/july97/>), and the forthcoming issue of *Archives and Museum Informatics: The Cultural Heritage Quarterly*, Vol. 11, No. 3, 1997.
- ³ 'Potentially involved' because professionally we still lack the capacity to mount effective records risk management cases relating to particular functions. Until we can do this our arguments on risk grounds will be dismissed by organisations chiefly concerned with the 'bottom-line'.
- ⁴ Reflected by some of the participants at the Electronic Records Research Working Meeting (See D. Bearman and J. Trant, *op. cit.*)
- ⁵ The terminology 'record core' obviously echoes the Dublin Core, a specification of metadata to provide a standard way of describing a wide range of 'document-like information objects' that would improve discovery of information resources available on the World Wide Web. The Warwick Framework, which complements the Dublin Core, proposes a container architecture for metadata packages of various types. The minimalist vs structuralist debate in the Dublin Core community is of interest in determining what metadata strategy might be adopted for recordkeeping. For a report on the 4th Dublin Core workshop, see 'The 4th Dublin Core Metadata Workshop Report' by Stuart Weibel, Renato Iannella and Warwick Cathro, *D-Lib Magazine*, June 1997 (<http://www.dlib.org/dlib/june97/metadata/06weibel.html>), and for up-dates on developments following the most recent meeting in Helsinki in October 1997, visit the Dublin Core site, (http://purl.oclc.org/metadata/dublin_core/).
- ⁶ The absence of international and national archival descriptive standards, such as ISAD(G) and RAD, from this list is deliberate as they are designed to achieve different objectives: that of describing a record or collections of records upon receipt into an archival collection or repository. The models discussed are premised upon attribution of metadata at time of creation of the record, a substantially different approach.
- ⁷ In Frank Upward's terms it is a 'generative document ... a starting point from which other things can flow', Teaching Notes to Monash University's Master of Information Management subject LAR 5640, Business Records Management, 30.9.1997. An example of the type of developments generated by the Standard is the Office of Government Information Technology's Records Management System Working Group's *Request for Proposal*, 8.10.1996 (<http://www.ogit.gov.au>). Section 3, Specific Requirements of the Records Management System, extends the Registration metadata specified in the AS 4390 *Australian Standard on Records Management*. See particularly section 3.3 Registration.

- ⁸ Traditional printed file covers often identify the agency as a banner heading, thus carrying this metadata.
- ⁹ 'High level metadata (what I call 'archival data') is needed to manage and comprehend records in any domain which exists outside of that in which the record originates.' Chris Hurley, 'Standards, Standardisation and Documentation' in *Archives at the Centre. Proceedings of the 1996 Conference*, Australian Society of Archivists, Canberra, 1997, p 64.
- ¹⁰ Chris Hurley: Presentation Script for his paper 'Standards, Standardisation and Documentation', presented to the ASA Conference Archives at the Centre 1996, and posted to the aus- archivists listserv on 29.5.1996. Archives of the listserv are available at (<http://www.asap.unimelb.edu.au/asa/aus-archivists/maillist.html>).
- ¹¹ UBC elements of archival bond and classification code; AS 4390 elements associated with links to related records documenting the same sequence of business activity and business system from which the record was captured. (See Table 1.)
- ¹² Much of the vision of how the recordkeeping processes would work in an item level environment is contained in D. Bearman, 'Item Level Control and Electronic Recordkeeping', *Archives and Museum Informatics*, Vol. 10 No. 3, 1996, and in draft form at (<http://www.sis.pitt.edu/~nhprc/item-lvl.html>).
- ¹³ This type of thinking is already present in the world of document management.
- The current stage of evolution is just now emerging. It's the stage where documents really get smart – they get active – and these documents often live on the Net and are dynamically generated and contain links and intelligence to many other pages and sites. In this world, documents are often transient things – they get personalised to individual users, they hardly ever see paper because that tends to make them less useful. These documents have a lot of intelligence – they are containers of content, links and programs – so they are not merely passive agents in a network, they can really be active citizens that can direct their own activity within the context of a DM system ... The goal of document management is no longer to manage documents. It is to teach documents how to manage themselves!
- Larry Bohn, 'Enterprise Document Management: the Next Generation', paper delivered at AIIM '97, downloaded from: (<http://www.pcdocs.com/>), April 1997.
- ¹⁴ This systematic attribution of disposal triggers can work either at an aggregate level associated with types of transactions or at an individual item level, depending on the nature of the business activity/transaction being undertaken.
- ¹⁵ For a more detailed exploration of these issues see Chris Hurley, 'Ambient Functions - Abandoned Children to Zoos', *Archivaria*, No. 40, Fall 1995.
- ¹⁶ *Ibid*, p.23.
- ¹⁷ Chris Hurley, Presentation Script, *op. cit.*

- ¹⁸ University of British Columbia, 'The Preservation of the Integrity of Electronic Records', Templates 5-8, downloaded from (<http://www.slais.ubc.ca/users/duranti/>) 7 July 1997.
- ¹⁹ Roman numerals relate to the numbering of the specifications contained in 'Metadata Specifications Derived from the Functional Requirements', downloaded from (<http://www.sis.pitt.edu/~nhprc/meta96.htm>) on 7 July 1997.
- ²⁰ Roman numerals relate to the numbering of the specifications contained in 'Metadata Specifications Derived from the Functional Requirements', downloaded from (<http://www.sis.pitt.edu/~nhprc/meta96.htm>) on 7 July 1997.