

**ADRI Workshop:  
Record Strategists and Archivists  
within Archival Institutions**

**Introduction to Digital Recordkeeping & Archiving**

**Manual and Course Material**

February 2010  
Prepared by State Records of South Australia

### Table of Contents

#### Day One Outline

Session 1	Introduction to Records and Information	6
Session 2	Recordkeeping in Government	11
Session 3	Digital Recordkeeping	14
Session 4	Digital Archiving	26
Session 5	Review and Preview	32

#### Day Two Outline

Session 6	Introduction and Review	33
Session 7	Standards for Digital Archiving	34
Session 8	Ingest	38
Session 9	Archival Storage	40
Session 10	Data Management	42
Session 11	Preservation Planning	44
Session 12	Access	46
Session 13	Administration	48
Session 14	Conclusion	50

© 2009 – 2010 Government of South Australia

This Course Material may be copied for use by South Australian Government Agencies and Local Government Authorities and by CAARA member institutions and for reasonable study or research purposes. No part of this Course Material may be reproduced or distributed for profit or gain or for any other purpose without the written permission of the Manager [Director] of State Records of South Australia.

## Workshop Rationale

This workshop aims to extend the collaborative nature inherent to the Australasian Digital Recordkeeping Initiative, to a learning environment. Collaboration is the underlying theme of the workshop, modelled in activities and workshop structure from beginning to end. The workshop also captures ADRI's collaborative approach by building on established processes, practices and initiatives within the various CAARA jurisdictions.

In developing this workshop, it was assumed that the audience are professionals – records strategists and archivists – who require an understanding of digital recordkeeping and archiving as a compliment to prior knowledge, skills and experience in a predominantly paper based environment.

The workshop aims to promote effective digital recordkeeping as conducive to effective digital archiving. In order to emphasize this continuity and facilitate learning, connections are made throughout the workshop to the corresponding functional entities in both these fields.

In November 2006, The CAARA Colloquium project identified the following important requirements for digital archiving in Australia:

- Prioritise elements needed to meet each of the functional entities of the OAIS Model
- Analyse existing paper archive controls and metadata used, and map these to the digital environment

These requirements have been incorporated into collaborative workshop activities. However, the option exists for the collaboration to continue beyond the workshop with workplace activities designed to assist CAARA member agencies. This would comply with an ADRI desire to work with a limited resource base.

# Introduction

## Learning Objectives

### Day 1

1. Define the terms “digital records”, “digital recordkeeping” and “digital archiving”
2. Identify the principles of recordkeeping within a government environment and discuss the issues relevant to both paper and digital recordkeeping.
3. Discuss current processes of archival management
4. Discuss the agency requirements needed for effective digital record keeping and how this will contribute to effective digital archiving

### Day 2

1. Define a model for digital archive development and the elements that contribute to its functional entities
2. Identify and prioritise elements of a digital archiving model for functional application
3. Use a functional model to develop solutions to archiving digital records in a contemporary archival environment

**What are the questions/queries you have that you would like to address over the two days?**

## Session 1:

### Introduction to Records and Information

#### Introduction

Within records and archival management there are particular standards and legislative requirements that need to be adhered to whether the record be of paper or digital format. Some of these requirements are legal documents that need to be followed and practiced by all agencies. Other standards are used to measure the effective implementation of recordkeeping procedures and are suggested as “best practice”. The various policies, standards and legislations have been developed to ensure that the requirements and accountabilities involved in records management are followed and practiced by all individuals within government. This legislation and standards will vary between jurisdictions.

#### What is the difference between information and records?

It is important to distinguish between records and other types of information, as there are different requirements for the management and retention of records. The distinction between records and other types of information is that records provide evidence of business activities.

#### Key Characteristics of a Record (AS ISO 15489)

The Australian Standard AS ISO 15489 explains these characteristics as:

##### *General*

A record should correctly reflect what was communicated or decided or what action was taken. It should be able to support the needs of the business to which it relates and be used for accountability purposes.

As well as the content, the record should contain, or be persistently linked to, or associated with, the metadata necessary to document a transaction, as follows:

- a) the structure of a record, that is, its format and the relationships between the elements comprising the record, should remain intact;
- b) the business context in which the record was created, received and used should be apparent in the record (including the business process of which the transaction is part, the date and time of the transaction and the participants in the transaction);
- c) the links between documents, held separately but combining to make up a record, should be present.

### *Authenticity*

An authentic record is one that can be proven to be what it purports and that it can be proven to have been created or sent by the person purporting to have created or sent it.

This means that organisations must have appropriate governance of its records and information management practices and systems, to ensure that it can prove a records authenticity through appropriate systems, policies, procedures and practices that ensure all activities on records (eg creation, maintenance, access, disposal) are controlled and protected against unauthorised alteration or deletion.

### *Reliability*

Reliable records are those that accurately reflect the business that is documented within them and can be trusted and depended upon for subsequent business.

### *Integrity*

Records that can be proven to be protected against unauthorised alteration (ie held inviolate), and have any authorised amendments fully acknowledged and detailed as such through appropriate and adequate metadata are considered to have retained their integrity.

### *Useability*

For records to be *useable* they must be able to be located and retrieved and understood in the broader context of the broader business. That is, they must be adequately connected to the business activity that generated the record and contain appropriate linkages to the broader business for the reader to fully understand the business that initially generated the records and any subsequent, associated activities.

## **Key Principles of Records Management (AS ISO 15489)**

Records are created, received and used in the conduct of business. To ensure compliance with the legislative and regulatory environment and provide necessary evidence, business should develop and maintain a records management program that includes:

- Determining what records must be created
- Deciding in what structure the records should be created and captured
- Establishing metadata requirements
- Determining retrieval requirements
- Assessing risks of not having records
- Preservation of records for access
- Compliance with legislative and regulatory requirements
- Maintenance of records in a safe and secure environment
- Appropriate retention of records
- Continuous process improvement

## Discussion

Discuss the key principles and characteristics of records as outlined.  
Discussion should focus on:

1. Do these principles and characteristics differ between paper and electronic records?
2. What impact do these characteristics have on an archival authority in relation to the management of its archival collection?
3. Are there any other characteristics of a record?

## Information and Knowledge Management

There is a significant move towards Information Management and Knowledge Management as disciplines.

To fully understand what Information and Knowledge Management are we first need to have some understanding of what 'Information' and 'Knowledge' is as concepts and how these interrelate to on another. To do this we are going to look at the model of:

To understand the difference between information and records a little further we will look at the model of:

**Data            Information            Knowledge**

### What is Data?

Data can be defined as symbols or unprocessed information that does not have any meaning by itself. Data is generally raw statistics and facts and is a prerequisite to information. Often data is stored in spreadsheets and databases, a collection of discrete objects.

Example: 25 degrees

### What is Information?

Information can be described as data that has meaning and context attached to it. In other words, information is data that has been processed in some way so we derive meaning from it. Often information answers Who, What, When and Where type questions for us. Example: 25 degrees Celsius temperature on Saturday.

In terms of Records Management, information only becomes a record when the "information created, received and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business" (ISO15489)

So, records always contain information however the opposite is not always true, information by itself does not always constitute a record.

## **What is Knowledge?**

Philosophers have debated what knowledge is for many hundreds of years dating back to Socrates (a Greek philosopher) who believed that knowledge was a virtue and the more you knew about yourself the better type of person you would be. There are many different attempts to define knowledge and many schools of thought as to what knowledge is depending if you are taking a psychological, religious or information profession background, which is the definition, we will be focusing on.

Knowledge can be seen as a subset of Information and is a result of learning. Knowledge occurs when information becomes meaningful to us and the intent of gathering the information becomes useful. Individuals have argued that knowledge often answers the 'Why are we doing this' and 'How are we doing this' type questions.

Example: 25 degrees Celsius temperature on Saturday, we will go for a walk on the beach

The Australian Standard on Knowledge Management (the AS 5037) argues that there is no single agreed definition of knowledge and describes knowledge as being multi faceted and can be:

- something residing in someone's mind (it may be highly personal and included aspects of culture or 'ways of doing things')
- it can be recorded as information in a document, image, film clip or some other medium
- it can be considered as a component of an organisations asset base

It is generally argued that there are two types of knowledge, tacit knowledge and explicit knowledge.

### **Activity: BBC – Science and Nature: Human Body and Mind**

<http://www.bbc.co.uk/science/humanbody/mind/surveys/disgust/index.shtml>

**Tacit knowledge**

Represents knowledge that is stored in someone's mind, personal knowledge and memory. An example given on Wikipedia in relation to tacit knowledge is the ability for one to recognize his or her's mother's face. In a matrix of 1000 similar faces, the son or daughter would almost always recognize the mother's face. However, he or she would have an extremely difficult time instructing a stranger how to identify the proper face. In this sense the son or daughter knows the mothers face in a way he or she cannot tell to someone who does not know the face personally.

It is much harder to measure tacit knowledge for this reason and it is more difficult to communicate to others. In an organisation tacit knowledge may be represented by only one person knowing how to complete a certain job or task, the way they do a particular thing is in their head.

A few ways we can pass on tacit knowledge is learning through observation, being mentored by or questioning/interviewing experts.

**Explicit knowledge**

Represents knowledge that is formal or recorded and can be easily communicated to others. Explicit knowledge is often represented within an agency through manuals, reports, documents, procedures, standards, books, documents etc. Often these are stored or captured electronically on various forms of media.

## Session 2: Recordkeeping in Government

### Back in the day...

That there are people in government in this city, in every building, on every second floor, deleting emails because they do not realise they are records, is indicative of the times we are in. Recordkeeping used to be a struggle about creating records, now it's not only a struggle with creating it is a struggle with capturing, defining what type of record it is and deciding what the hell we do with it!

It is no overstatement to claim that record keepers make history. Those that manage these records are the curators of tomorrow's history.

In the olden days it used to be quite simple...generally there were a whole heap of letters sent or received, written quite elegantly and with the Queens English of course, and these were registered in some equally beautiful calf skin bound leather volumes manufactured by a character from Dickens.

Limited types of information in a limited format, managed with methods that were almost as immutable as stone.

The letters, and whatever else had to be documented, came on paper. There was lots of paper, as there was very little else...maybe vellum, leather, but certainly lots of paper.

The records created in this relatively stable format had a shelf life bearing pretty good odds. And even better, you didn't need anything else other than a magnifying glass, at the worst of times, to be able to read them. One hundred, two hundred years later, that's probably still all you need.

It was by pure chance that most of what survives today has done so. No legislation or standards to govern how records were created and maintained and disposed of. Neither was there the significant demand from the public to access this material. Was their litigation of who said what to whom and when? Was the media scrutinizing the actions and behaviours of all public servants?

### **Activity – Work in small groups to discuss and present on the following:**

Do you agree with the paragraph above, *especially the statement that “It was by pure chance that most of what survives today has done so”*

Take this example from South Australia. In the 1920s, records in the South Australian Archives could only be viewed by “bona fide students bearing a letter of introduction from the Vice-Chancellor of Adelaide University.” And these records tended to be those relating to the high offices of government. The records relating to the lives of the masses were left untouched (SRSA Ancestors in Archives 2000:x)

The greatest threat to paper records was fire and flood.

This was the case well into the 20<sup>th</sup> century.

**Activity – Insert State appropriate record example and context behind the record chosen including metadata eg. Description, date, format etc...**

*Participants to work in small groups to discuss and present on the following:*

What records may have been created?

What issues surrounded their retention?

What are the preservation issues?

How are these records, including the proclamation, accessed?

### **Where are we now?**

Within the past 2 decades, the story has changed dramatically. A significant amount of us have broadband on in our homes, and an even larger amount has access to the Internet on our phones (even though half of us don't know how to use it!). The point is, at the flick of a switch we can be inundated with information from a variety of sources and in a variety of formats.

Information is at our fingertips, we want it all and we want it now. Governments have a responsibility to service communities and as technology changes so do the demands of our clients. It is now not enough to provide information via an office or over the phone, transactions between members of the public and governments are conducted online, in an environment where a traditional paper record may never be created, let alone captured and disposed of.

In addition to this and in stark contrast to the 1920s, it's the records about the everyday person that attracts the attention, with the vast majority of archive users being family historians (SRSA Ancestors in Archives 2000:x).

Paper is now like some kind of sacred cow that should not be used on any day except Sunday, and it is the digital, electronic record that has come to the fore. But we need more than a magnifying glass for these, and we can't immediately guarantee that these records will be accessible on the shelf one hundred, two hundred years later.

**Activity – Insert State appropriate record example and context behind the record chosen including metadata eg. Description, date, format etc...**

*Participants to work in small groups to discuss and present on the following:*

What issues surround the records retention?

What preservation issues are there?

How is this record accessed?

Things change. The topic of a relatively recent RMAA conference was “evolution”. In training programs across Australia, records management professionals learn about implementing RM programs that are able to take the punches from those creatures of change that appear on the horizon, or are waiting for them at the top of the plateau they envisage as being the end of their struggles. Records management is constantly evolving and it has a flow on effect for all other professionals within the realm i.e. archivists etc. One of those creatures has been electronic records. We would all do well to learn the ways of our government recordkeepers, to understand the sorts of history that will be coming our way as archivists.

## Session 3:

### Digital Recordkeeping

*The information contained below, unless cited otherwise, has been drawn from the UniSA BIM Course material for Digital Recordkeeping, developed by Judith Ellis, 2008*

#### What are digital records?

##### **Electronic Record**

A record created, communicated and/or maintained by means of electronic equipment. Although this term can refer to analogue materials (eg. videotapes), it generally refers to records held in digital form on magnetic or optical computer storage media.

*(Adapted from Standards Australia AS 4390, Part 1, Clause 4.13; Kennedy, J, and Schauder, C, Records Management: A Guide to Corporate Recordkeeping, 2nd edition, Longmans, Melbourne, 1988, p. 293.)*

##### **Digital Record**

A record created and/or maintained by means of digital computer technology. Includes records that are 'born digital' or have undergone conversion from a non-digital format. Digital records are a subset of electronic records.

*(National Archives of Australia, Guidelines for Implementing the Functional Specifications for Recordkeeping Functionality in Business Information Systems Software Exposure draft, 2006, p.107, <http://www.naa.gov.au/records-management/publications/BIS-guidelines.aspx>)*

The propensity for Government to build information systems based on computers and computer networks poses a formidable challenge to agencies charged with custodianship of the State's memory. The knowledge and tools required to manage digital memory for persistence are very different from traditional records. Digital records exist on fragile media and cannot be seen. Combinations of hardware and software are required for digital records to be understood. As hardware and software become obsolete, the essential conditions for intelligibility may no longer be satisfied resulting in loss of digital memory. <http://www.sro.wa.gov.au/dri/dri.asp>

Prior to the 1990s, few Australian recordkeepers had seriously addressed electronic records... The records tradition in Australia is one which has been dominated by the inheritance of registry systems from late eighteenth- and nineteenth-century England. As a colony we were subject to systems of

bureaucratic control designed to manage at a distance. The emphases of the registry system, with its pre-action and process control implications, have served us well in the transition to electronic recordkeeping. Records management in Australia from the 1950s was not limited to minding and managing the non-current files after action was completed - the prevailing American notion of records management. When computers did hit records management, they first appeared in the mid 1980s as systems aimed at automating these registry controls... The extraordinary level of recordkeeping activity in Australia in the past two years has been aimed at establishing frameworks. It explicitly recognises that new approaches are needed to deal with the reality of the electronic workplace  
Barbara Reed, Electronic records management in Australia, *Records Management Journal, UK, Vol 7 No 3, December 1997*, [http://www.recordkeeping.com.au/pub\\_ermaust.html](http://www.recordkeeping.com.au/pub_ermaust.html)

## **Web 2.0 Technologies**

The term was coined to represent a shift in the use of the internet. The so-called "Web 1.0" environment, a title designated retrospectively, was marked by the publication of information to the internet, with little or no user control or input. For example, a website with content published by the creating agency. Web 2.0 provided the technology for users to not only comment on the information, but to control it, and to carry out mass interaction on a global scale. This is particularly evident in such technologies as blogs, wikis, social networking systems (like twitter, facebook and ning) and instant messaging.

Web 2.0 applications, and combinations of them known as "mash-ups", are increasingly finding a use in government business. Consequently, it is likely records will be created as a result of using these applications. Information and records management professionals face a challenge when it comes to managing Web 2.0 as a medium for official records, largely a result of standards and frameworks being absent from a rapidly evolving domain. In order to better cope with these challenges, information management professionals need to integrate web 2.0 applications into their records and information management programs

<http://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/guidance/guidelines/Guideline-24-Records-management-and-web-2.0>

State Records NSW Guideline 24 Records Management and Web 2.0

## Digital Records and The Records Continuum

A Records Continuum Model was developed by Frank Upward in 1996 of the *Records Continuum Research Group* at Monash University. It has been explained and further discussed by Frank Upward in his papers, *Structuring the Records Continuum - Part One: Post Custodial Principles and Properties* and *Structuring the Records Continuum Part Two: Structuration Theory and Recordkeeping*

In 1996 the records continuum was described as being “the whole extent of a record’s existence. Refers to a consistent and coherent regime of management processes from the time of the creation of records (and before creation, in the design of recordkeeping systems), through to the preservation and use of records as archives.” (AS4390 Australian Standard for Records Management, 1996, superseded by AS ISO15489 Australian Standard for Records Management).

Prior to recognition and acceptance of the records continuum concept, the records “life-cycle” was used to describe the stages of the “life” of records. The “life-cycle” implied a ‘beginning and end’ or ‘birth and death’. The ‘beginning’ or ‘birth’ being record creation and the ‘ending’ or ‘death’ being disposal. The flaws of the “life-cycle” model include:

- it separates the activities of Records Managers and Archivists
- it does not account for the stages prior to records creation, eg, the specification and design of recordkeeping systems
- it implies there is no ongoing value in records after disposal
- no consideration is given to requirements for preservation and access to records as archives
- it is records-centered, rather than business or process centered
- it is concerned with records as physical entities
- it is locked into custodial roles and strategies
- it is operationally focused and concerned with records management tasks (things done to the records in fixed stages) rather than the integration of business and recordkeeping processes.

The records continuum, on the other hand, has provided Australian records managers and archivists with a way of thinking about the integration of recordkeeping and archiving processes.

A continuum-based approach suggests integrated time-space dimensions. Records are 'fixed' in time and space from the time of their creation, but recordkeeping regimes carry them forward and enable their use for multiple purposes by delivering them to people living in different times and spaces. (Sue McKemmish , *Yesterday, Today and Tomorrow: A Continuum of Responsibility*, 1997)

Records have multiple purposes, or in terms of the more traditional life-cycle view, we could say that records can have multiple lives. They have multiple purposes in terms of their value to an individual, an organisation or society. And that value could be at different times, it could be immediate value to an organisation that created the record, and could be in 20 years time to an economic historian writing a paper on a particular topic.

Records are “vehicles of communication and interaction, facilitators of decision-making, enablers of continuity, consistency and effectiveness in human action, memory stores, repositories of experience, evidence of rights and obligations. ....their utility, their recordness, is bound up with their evidentiary qualities defined in terms of their transactional and contextual nature.” (Sue McKemmish, *Archives: Recordkeeping in Society*, p.15, 2005 – see references at the end.)

In her paper “Yesterday, Today and Tomorrow: A Continuum of Responsibility”, Sue McKemmish describes how record continuum thinking and practice is underpinned by a concept of records of continuing value. She states that records continuum thinking is concerned about ideas about the role of recordkeeping:

- Records play a role in governance, in particular regulating relationships between people and organisations, and as instruments of power and authority.
- There is the nexus between recordkeeping and accountability in its broadest sense of accounting to each other for what people do to each other, encompassing corporate, social, cultural, and historical accountability.
- Recordkeeping plays a role in constituting corporate and collective memory, especially insofar as records capture experiential knowledge.
- Records play a role in that recordkeeping can be understood as a witness, providing evidence of both personal and collective identity.
- A final role is the way records function as sources of value-added information and can be exploited as assets, with new records being created in the process.

(McKemmish, S – *Yesterday, Today and Tomorrow: A Continuum of Responsibility*, 1997)

From this understanding of the special nature of records and archives (ie. their evidential quality), we can, in turn, see it is the responsibility of the recordkeeping and archival profession/s to “build and manage frameworks and systems which assure the preservation and accessibility of accurate, complete, reliable, and authentic records and archives...The accountability of the recordkeeping profession and archival institutions is therefore a critical issue in democratic societies”. (Sue McKemmish, *ibid*, p.16)

The records continuum approach strongly links the mission of recordkeeping professionals in Australian society to corporate, democratic and historical accountability, and more broadly to the role of recordkeeping in society (the sociology of recordkeeping).

It emphasises the steering role of recordkeeping professionals as:

- recordkeeping policy makers
- standard setters
- designers of recordkeeping systems and implementation strategies
- consultants
- educators/trainers
- advocates
- auditors.

The records continuum aims to redefine and reinvent recordkeeping practice and processes, including appraisal, control, storage and access. Recordkeeping needs and activities are layered upon each other rather than occurring in stages as they are depicted in a linear life-cycle approach.

## **Records Systems**

“An information system which captures, manages and provides access to records through time.” (AS ISO 15489-2002, Records Management).

Managing records as evidence of social and organisational activity involves incorporating them into records systems that can:

- carry them forward with their 'fixed' content
- re-present their structure or documentary form
- maintain sufficient contextual links to preserve their meaning through time.

Just as records constitute a special type of document, so records systems are a special type of information system. This was first clearly articulated by David Bearman (in "Record-Keeping Systems", *Archivaria* 36 (Autumn 1993): 16-37). Information systems are normally designed to keep information that is:

- timely
- non-redundant
- manipulable.

To act as records systems, they need to contain information that is:

- time-bound
- redundant (ie. information able to be omitted without loss of meaning or function)
- non-manipulable.

Electronic document/records management systems (EDRMS) are a type of information system whose primary purpose is the capture and management of digital documents and records. EDRMS are systems designed specifically to manage the creation, use, maintenance and disposal of digital documents and records for the purposes of providing evidence of business activities.

EDRMS are distinguished from other information systems by

their ability to:

- maintain appropriate contextual information and metadata, and links between records to enable their identification, support their value as evidence and provide access to them over time;
- enable the application of records management processes, such as classification, registration, search and retrieval, preservation and disposal; and
- apply controls to records, such as access and security controls, to preserve their content and secure their integrity.

*(Above two paragraphs from: National Archives of Australia, Guidelines for Implementing the Functional Specifications for Recordkeeping Functionality in Business Information Systems Software Exposure draft, 2006, p.37)*

EDRMS present one of the best method for maintaining digital records over time, as they provide digital records with the necessary context and safeguards to assist in their long-term preservation.

Note some other types of information management systems are also starting to include the capability of managing records, for example:

- Web Content Management Systems
- Enterprise Content Management Systems
- Email Management Systems
- Workflow Systems
- Digital Asset Management Systems

The main purpose of a recordkeeping system is to capture, manage and maintain recordkeeping processes associated with records.

## **Discussion**

What processes are undertaken to achieve the capture, management and maintenance of records within a system?

Points covered should include:

- Capture
- Registration
- Classification
- Access
- Activity tracking
- Location tracking
- Workflow
- Re-Use
- Links to other information
- Disposal
- Status or disposition of the record (active, inactive, archived, destroyed)
- Reporting

## **Electronic Document and Records Management System**

An EDRM system is designed to manage electronic content, documents and records and support four key functions:

- input (creation/capture)
- management (content, documents, records)
- collaboration/process management
- output/delivery.

An EDRM system is a combination of an active electronic content and document management system and an electronic records management system.

### **Creation Capture and Control**

EDRM systems can be configured to meet the needs and the business processes of an organisation. The types of records being created and managed by the organisation, the organisation structure and its business activities will determine what elements are included in the configuration of software. When an organisation purchases an EDRM system, there are many preliminary steps required prior to implementation, including determining how the system will look to the end user such as entry forms and dialog boxes.

The capture of records into a recordkeeping system is the first step in the management of an organisation's records.

The purpose of capturing records into a recordkeeping system is to:

- Create a relationship between records and their business context
- Link records to other information
- Enable ongoing access to records
- Provide records with reliability and authenticity

In the past capturing and managing text, graphics and images has been cumbersome especially as those in paper, microfilm, CD ROM format. High-end EDRM systems provide functionality previously provided in multiple systems, some of which were manual, these include the ability to:

- attach metadata to a document/content object;
- register each document/content object and log the document/content object into a repository;
- apply check out and check in procedures and version control facilities so document/content objects cannot be overwritten;
- apply document/content object level security so only authorised users can access and read the document/content object;
- provide a subset of users the ability to edit the document/content object;
- use a mix of structured index data and full text indexing to provide a flexible range of searching and retrieval options.

### **Activity**

Compare traditional methods of creation, capture and control with contemporary ones, e.g. in a digital environment. What are the differences and issues?

### **Access and Security**

In any organisation there are access restrictions applicable to information by both internal and external sources. The complexity of these restrictions is dependent on the business or personal sensitivity of the information. For example access to personnel records are restricted in all organisations. 'Product development' information of a manufacturing organisation would be subject to internal and external access restrictions to protect the intellectual property.

External access is controlled by policy and procedures and the assignment of responsibility for handling requests for access to designated staff. Internal access to physical or hardcopy information is controlled by the individuals who are responsible for a particular activity separately managing that information. They securely store the information and manage any internal access requests.

Access to information held in an enterprise-wide system, such as an EDRM system relies on the configuration and allocation of security regimes.

These security regimes should apply to documents and records at the time of creation, receipt and capture to ensure protection of any confidential, private, sensitive or intellectual property from inappropriate access, usage, disclosure or alteration.

Security levels are hierarchical by nature and can be applied as a default on Record Types, Locations, and Classification/File Plans or applied to individual records on a case-by-case basis.

Common components to control security are:

- Security levels –hierarchical in the level on control
- Security caveats – add a layer of access restrictions
- Access controls – can be applied to many objects and can control who can view or update records.

Additional layers of security or ‘access control’ for records can be designed for complex security requirements. Such a three-dimensional security model can be applied to:

- Records and other object types in the system – at the record type or individual record level – done through ‘access controls’
- Locations (people/roles)
- Classifications (levels of the classification scheme)

Access controls provide for the restriction of users ability to view or update metadata and records. Access controls can be applied to the various objects in an EDRM system and can be additional to, or used instead of Security Levels and Caveats.

Access Controls can be applied to the following objects and records:

- Record Types
- Classification Plan
- Thesaurus
- Retention Schedules
- Reports
- Locations
- Lookup Sets
- Records

A precursor to any output is searching. Once records are captured into a system they are subject to recordkeeping processes and business activity depending on their content and purpose. They may be accessed, relocated and have related information linked to them or may never be looked at again.

The reasons for searching:

- Reuse information
- Validate information
- Evidence for accountability
- Discovery Orders
- Evaluate programs or projects

- Reporting
- Planning

If records, or the information relating to records, are required then the system will be searched to retrieve the required information. It is vital that the information registered about a record is accurate and comprehensive to ensure records are not only valid and have meaning, but are able to be found.

EDRM systems provide simple and complex searching, by a variety of search methods.

### **Activity**

Compare traditional methods of providing access and security with contemporary methods in a digital environment. What are the differences and issues?

### **Metadata**

Recordkeeping metadata for records captured into an EDRMS provide context, content and structure of the records. It is imperative that the metadata be appropriately managed to ensure the evidential value of the records.

Recordkeeping metadata must be managed, maintained and retained for as long as the record is retained, in an open and enduring format. Even when records are removed from the system, the metadata associated with the records provides evidence of the existence of the records.

Recordkeeping metadata includes all standardised information that identifies, authenticates, describes, manages and makes accessible through time and space documents created in the context of social and business activity. Traditionally some of this metadata has been captured in records systems and some in archival control systems and finding aids and some of it has been present in the physical form, ordering, juxtaposition and location of records. Increasingly recordkeeping metadata is captured in workflow, document management and knowledge management systems, and it is essential to make what was before evident in the physicality of the record explicit in metadata.

When records move beyond the boundaries of the local domain in which they were created, or, as is the case in networked environments, they are created in the first place in a global rather than a local domain, this kind of metadata needs to be made explicit (i.e. captured and persistently linked to the record). This is essential so that users in the broader domain can uniquely identify, retrieve, and understand the meanings of the records.

In the hardcopy world this need typically arose when records were transferred to an archives repository. Broader contextual metadata was captured in archival control systems and finding aids rather than in current records systems. However, in the global and virtual domains of cyberspace, we need to make what was known before through physical ordering and location (custody) explicit in metadata captured with the record or persistently linked to it. We need to document fully the logical associations that derive from the role the records play in business processes and contexts and we need to make available to current records systems the broader contextual metadata traditionally found in archival systems and finding aids.

For example, information about the records' wider organisational context and relationships to high-level functions needs to be captured, as well as the narrower contextual metadata of the immediate business environment. The latter includes information about transactions, business processes, and the actors involved - that is captured in enterprise systems such as document management, workflow, and human resources applications.

Digital recordkeeping processes related to metadata include the following.

**Capture** - a deliberate action, which results in the registration of a record into a recordkeeping system. (AS 4390.1 General, clause 4) It is similar to 'registration'.

**Registration** – the act of giving a record a unique identifier on its entry into a system (AS ISO 15489-1). The purpose of registration is to provide evidence that a record has been created or captured in a recordkeeping system. It involves recording brief descriptive information about the record in a register and assigning the record a unique identifier.

**Indexing** – the process establishing access points to facilitate retrieval of records and/or information. In electronic systems registration and indexing are the same process, for example through the allocation of codes, locations and a system for retrieval.

**Classification** – the systematic identification and arrangement of business activities and/or records into categories according to logically structured conventions, methods, and procedural rules represented in a classification system. (AS ISO 15489-1)

**Controlled vocabulary** - further descriptive and control details can be attached to the record by using vocabulary controls such as a list of authorized headings or a thesaurus.

#### **Activity**

Compare traditional methods and types of metadata with contemporary methods in a digital environment. What are the differences and issues?

## Disposal

The requirement for organisations to manage the retention and destruction of records as prescribed by legislation and disposal schedules, based on legal, fiscal, historical, community and business requirements, was identified and described in previous lectures. The application of disposal schedules to records to determine their retention periods is undertaken within the EDRM system by Administrators. The Administrator is required to input each disposal entry or class into the online schedule. In most cases schedules are created within a specific template for approval by an authorising body in the various jurisdictions or within private sector organisations by senior executive or Board members. Most EDRM systems enable the importing of disposal schedules from specific formats which is the preferred method as there is room for human error when manually inputting information.

Each record class or record description requires a separate entry. Retention periods and disposal actions are created for each entry and are identified as Triggers.

### Activity

Compare traditional methods of undertaking disposal with contemporary methods in a digital environment. What are the differences and issues?

### Podcast – Records Management 2.0

[http://nuweb.northumbria.ac.uk/ceis\\_podcasts/2009/episode04.php](http://nuweb.northumbria.ac.uk/ceis_podcasts/2009/episode04.php)

- what impact is Web 2.0 having on the way organisations are keeping their records?
- are current records management practices and standards still adapted to the web 2.0 world with its increased volume, and pace of information exchange, increased diversity of systems and increased pace of technological change?
- what kind of record keeping would be suited to the web 2.0 world? Will the web 2.0 world result in organisations keeping records in a completely different kind of way?

## Session 4:

### Digital Archiving

#### What is digital archiving?

Digital archiving is the preservation and maintenance of digital records that have been either created in an electronic format or migrated into an electronic document.

#### Archival Principles

How does the permanent retention of digital records differ from hardcopy records?

For hardcopy records two important Archival Principles are:

- Provenance
- Original Order

Provenance and Original Order help the user of archives feel confident that what they are accessing are reliable, authentic records of functions and activities undertaken in the past.

#### *How do we establish Provenance?*

Through the collection of metadata at the time of accessioning the records, including:

- Name of the agency, organisation, person or family that created, maintained and used the records (creator)
- Classification terms (function/activity of the record)
- Series (links to other records created for the same purpose)
- Through an agency/creator description that details the administrative history of an agency/creator
- Through a Series description that provides the context for the creation, maintenance and use of the records, and describes relationships to other series of records created by the same agency, organisation or person.

#### *How do we establish Original Order?*

Through research and description of how the records were:

- collected (eg by subject, by author, by date etc.),
- stored (eg as a file or docket, as a batch, as an individual document etc.) and/or
- indexed (eg by file number, alphabetically by subject or author, etc.)

In addition to this, we must also address the concepts of:

- Authenticity
- Reliability, and
- Integrity.

An authentic record is one that is capable of being proved to be what it purports to be (i.e. the content is what it appears to be, it was created by the person who appears to have created it, and it was created at the time it appears to have been created).

A reliable record is one that contains a full and reliable representation of the facts which the record documents.

Integrity refers to the record being complete and without unauthorised alterations.

*So, what is different about archiving digital records?*

The challenge in preserving electronic records is ensuring that the systems that manage the electronic records hold sufficient metadata and implement suitable processes to ensure the long-term retention of context, authenticity, reliability, and integrity.

*(Archival Principles is taken from the Business Information Management program, Digital Recordkeeping, Lecture 13 by Judith Ellis 2008)*

## **What is happening around the world?**

*The National Archives, United Kingdom*

The National Archives (TNA) collects records as defined by the Public Records Act 1958. Essentially the archive collects records of central government, the courts and public enquiries. More well known paper material within the archive include the Domesday Book, census data and Cabinet Office Minutes.

TNA holds the bulk of its digital records in its Digital Archive. This material includes digital objects from a wide variety of types – documents, images, databases, emails, video audio etc. Other born digital material is held in a Web Archive (for web sites, naturally) and The National Digital Archive of Datasets which holds datasets from central government departments (generally, survey or census type materials which are in a raw form and require substantial supporting information for interpretation).

TNA's Digital archive does not prescribe a particular file format for collection of digital objects. Consequently, there is potentially a very wide range of digital object types to be transferred to TNA by the government agency that created them. However, TNA does provide guidance on format selection for preservation planning.

Metadata accompanying a submission are strictly controlled and requirements are built around the type of descriptive information provided. This is all undertaken in collaboration between Records Officers within the government agencies and staff from TNA. All this submission information goes through a process of validation at TNA and is sometimes further enhanced, along with any technical information.

Archival storage within TNA for digital objects is a service with limited responsibilities that is easily resolved with the help of technology. TNA relies on advanced technology for their storage management, but also use a policy of multiple copies for objects to reduce risk of data loss. Describing the archival collection from a digital perspective was a matter of building onto a long tradition of archival description applied to traditional media. The administration of the digital archive is spread across departments and units and responsibility is shared amongst employees.

Under the banner of preservation planning, TNA is active in setting preservation standards, building technologies to support migration planning and monitoring technology and file format identification

Users of TNA digital archives gain access via reading rooms or from the internet.

#### *PROV – Public Records Office of Victoria*

The Public Record Office of Victoria is the state's archival office holding records that date back to the mid 1830s. The office contains both traditional paper based records and also electronic records. The digital archive at PROV will only accept records that have been maintained in an electronic recordkeeping system that is compliant with the Victorian Electronic Records Strategy (VERS).

VERS originated in 1995 when the Victorian State Government together with industry and academia tried to find a practical solution to dealing with electronic records. The initial standard was launched in 2000 with the current updated version released in 2003. VERS provides a solution to the creation, capture, and preservation of electronic records. There is a specific strategy that has been adopted by VERS, which specifies the standard format for electronic records which:

- Is generic but extensible, so that it will work in conjunction with existing recordkeeping and business practices
- Ensures that all records are stored in a documented format, to enable viewing of records in the future, regardless of the system that created them
- Specifies methods to automate the capture of records from the desktop and business systems
- Specifies ways and forms in which to capture information about records and encapsulate this with the records to ensure that records in the future will be understood in context
- Details a method for securing records so that any changes are detectable

(<http://www.prov.vic.gov.au/vers/vers/strategy.htm>) as sourced from the VERS website

There are various benefits associated with the use of the VERS framework including, but not limited to, the following:

- Significantly improved accessibility to records over time and distance
- Reduced paper record handling and storage
- Increased discovery and reuse of records

Prior to the transfer of records PROV will test the Agency's system to ensure that it is VERS compliant and it meets the requirements of the VERS standard.

A manifest of the records that the agency wishes to transfer accompanied with the relevant documentation is emailed to PROV. All permanent electronic records need to be transferred to PROV as VEOs (VERS Encapsulated Objects). A VEO is an electronic record that has been formatted to the specifications and standards as outlined by PROV and VERS.

The stages of transferring the electronic records are:

1. Transfer the VEOs to PROV either by using a secure site on the internet or physical media.
2. Once the VEOs are received they are validated against the manifest and the validation rules, including virus checks, that PROV has in place
3. If any problems are detected at stage 2 then PROV has the right to stop the transfer and request the manifest to be retransmitted
4. The VEOs sit in quarantine for seven days and checked for viruses for a second time.
5. The VEOs have been cleared for transfer into the Digital Archive
6. PROV distributes a "custody report" to the transferring Agency stating that PROV formally accepts custody of the VEOs
7. Agency is granted permission from PROV to destroy any remaining copies of the VEOs that they have.

Further information is available at

[http://www.prov.vic.gov.au/records/transfer/electronic\\_records.asp](http://www.prov.vic.gov.au/records/transfer/electronic_records.asp)

#### *NAA – National Archives of Australia*

The National Archives of Australia is a federal government agency that houses a collection of records ranging from paper files to photographs to radio and speech recordings and also including digital records. The NAA has its main office located in Canberra with additional offices located in each state's capital and Darwin.

NAA's digital archive was established in 2006 and has been designed to:

- Encourage use by anyone needing to access digital records over a period of time
- Convert the files formats into "open formats" which allows a greater potential lifespan in the future
- Operate across various platforms

XENA (XML Electronic Normalising of Archives) software is what NAA use to their digital preservation process. This software converts digital format from their original format into an open format that can be used by NAA.

The stages of transferring electronic records into the NAA digital archive is:

1. Digital records in the form of CD, DVD or USB are received by NAA
2. The record is connected to a quarantine facility where the manifest is checked to ensure that the correct files have been received.
3. The contents of the record are copied onto a USB and are checked to ensure the integrity and authenticity of the original data.
4. The information copied onto the USB is placed in quarantine for 28 days.
5. Antivirus checks are conducted throughout the 28 day quarantine period
6. After the 28 days the USB is scanned again to detect any viruses
7. The USB is then connected to the “preservation facility” and the checksums are checked
8. The information is then processed to produce 2 versions of the record, one as an encoded version that contains XML metadata header and footer while the other is a copy in an open format
9. The new copies are placed onto another USB and connected to the digital archive
10. The data is copied onto RAID storage in the repository
11. Digital Preservation Recorder (DPR) software captures descriptive and preservation metadata
12. Data can then be copied out of the archive for access
13. There is continual monitoring of the record to ensure that the copies have not changed or altered. Any unplanned changes can be detected throughout the life of the record

The process that NAA uses has a number of checks to make sure that the record is not corrupted in any way and maintains it's authenticity and integrity while the transferred records are undergoing the transfer process into the Digital Archive.

### Activity – Comparison of PROV and NAA approach

- Divide participants into two smaller groups.
- One group is to list the pros and cons of PROV while the other group is to list the pros and cons of NAA (approx 5 –8 minutes)
- Write answers on whiteboard to share with the greater group and discuss (up to 10 minutes)

Differences relating to PROV & NAA Digital Archives

**PROV** – Requires agencies to maintain records as VERS compliant

**NAA** – Doesn't require agencies to have certain systems in place. If an agency requires tools to create a transfer manifest for transfer of electronic records then they will supply the necessary tools

**PROV** – Requires records that are going to be transmitted in a already particular format (VEOs). It is the transferring agency's responsibility to have tools that are able to process the permanent records that are compliant with the VERS standard

**NAA** – Receives the records in the original format from the transferring agency and NAA undertakes the conversion itself. They have more control over the metadata information that they input onto the digital record

Further information available from the following websites:

<http://www.naa.gov.au/>

<http://www.naa.gov.au/records-management/index.aspx>

National Archives of Australia (NAA)

<http://www.prov.vic.gov.au/>

<http://www.prov.vic.gov.au/about/>

Public Records Of Victoria (PROV)

<http://www.prov.vic.gov.au/vers/standard/>

<http://www.prov.vic.gov.au/vers/vers/>

Victorian Electronic Records Strategy (VERS)

## Session 5:

### Review and Preview

#### **Review of Day One**

Today's sessions have introduced us to the world of records management, and particularly the complexities and issues involved with managing records in a digital environment.

We have been introduced, albeit briefly, to the concept of a virtual archive.

Questions and comments.

#### **Preview of Day Two**

Tomorrow's sessions will concentrate on the virtual archive. Discussion will be based on the Open Archival Information System (OAIS) model, which has been endorsed and adopted by ADRI. Participants will work through the various stages and entities involved in the OAIS model.

Questions and comments.

## Session 6:

### Introduction to Day Two

#### Day One Refresher

Yesterday we discussed recordkeeping and the impact this has on archiving and the management of an archival collection, with particular reference made to the impact and issues raised through the dramatic increase in digital records.

We talked about how an agencies ability to manage their records appropriately will greatly improve the likelihood that these records will be suitable for management and preservation within an archive.

We also looked at how a number of archival institutions are already well progressed along the road to managing archival collections consisting of both paper and digital objects.

#### Outline of Day Two

Today's sessions will focus on expanding your understanding of how a virtual archive operates and subsequently the impact this has on people, technology and processes.

Today we will look closely at the OAIS model, the fundamental elements that it includes and how it differs from a traditional archival environment.

Much of the material we will discuss during today's sessions is referenced from the following documents:

*CCSDS Reference Model for an Open Archival Information System*  
<http://public.ccsds.org/publications/archive/650x0b1.pdf>

*ADRI Model Plan for an Archival Authority Implementing Digital Recordkeeping and Archiving*  
<http://www.adri.gov.au/model-plan.doc>

## Session 7:

### Standards for Digital Archiving

#### The Changing Environment

The archival environment has changed; no longer are archives “institutions” that collect manuscripts for purely research purposes. The information being retained and preserved by archives has changed, the format of this information has changed, the people accessing these objects is changing, even the general understanding of what constitutes an archive has changed.

As the Consultative Committee for Space Data Systems (CCSDS) state in their *Reference Model for an Open Archival Information System* (Blue Book, January 2008, Page 2-1):

*The term ‘archive’ has come to be used to refer to a wide variety of storage and preservation functions and systems. Traditional archives are understood as facilities or organizations that preserve records, originally generated by or for a government organization, institution, or corporation, for access by public or private communities. The archive accomplishes this task by taking ownership of the records, ensuring that they are understandable to the accessing community, and managing them so as to preserve their information content and authenticity. Historically, these records have been in such forms as books, papers, maps, photographs, and film, which can be read directly by humans, or read with the aid of simple optical magnification and scanning aids. The major focus for preserving this information has been to ensure that they are on media with long-term stability and that access to this media is carefully controlled.*

*The explosive growth of information in digital forms has posed a severe challenge not only for traditional archives and their information providers, but also for many other organizations in the government, commercial and non-profit sectors. These organizations are finding, or will find, that they need to take on the information preservation functions typically associated with traditional archives because digital information is easily lost or corrupted. The pace of technology evolution is causing some hardware and software systems to become obsolete in a matter of a few years, and these changes can put severe pressure on the ability of the related data structures or formats to continue effective representation of the full information desired. Because much of the supporting information necessary to preserve this information is more easily available or only available at the time when the original information is produced, these organizations need to be active participants in the long-term preservation*

*effort, and they need to follow the principles espoused in this OAIS reference model to ensure that the information can be preserved for the Long Term. Participation in these efforts will minimize the lifecycle costs and enable effective long-term preservation of the information.*

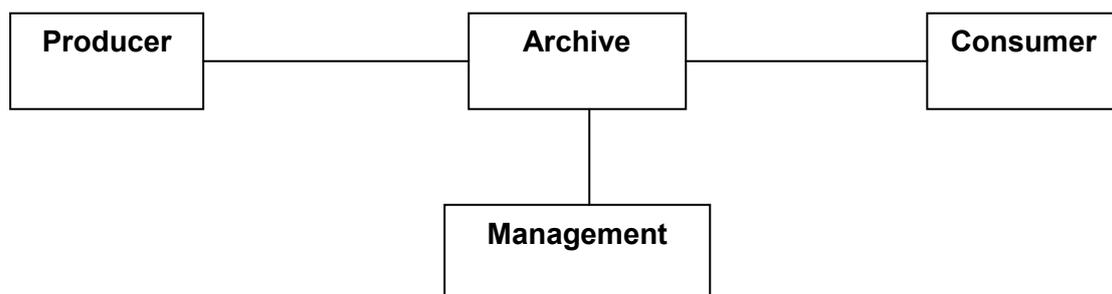
## **The OAIS Model**

The model for a virtual archive, established by the CCSDS, is a framework that specifically applies to organisations with a responsibility to make information available in the long term. It therefore applies to the “traditional” archive as well as an increasing number of other organisations who have a business need to ensure the long-term preservation of and access to, records.

The CCSDS state that:

*An OAIS archive is one that intends to preserve information for access and use by a Designated Community. It includes archives that have to keep up with steady input streams of information as well as those that experience primarily aperiodic inputs. It includes archives that provide a wide variety of sophisticated access services as well as those that support only the simplest types of requests.*

The following model demonstrates simply the environment surrounding an OAIS archive:



In this model the following definitions apply:

*Producer* is the role played by those persons, or client systems, which provide the information to be preserved.

*Management* is the role played by those who set overall OAIS policy as one component in a broader policy domain.

*Consumer* is the role played by those persons, or client systems, which interact with OAIS services to find and acquire preserved information of interest. A special class of Consumers is the Designated Community. The

Designated Community is the set of Consumers who should be able to understand the preserved information.

As you can see this approach is similar to how a traditional archive may view itself and some of its stakeholders.

Under the OAIS model the archive consists of a number of functions / functional entities. These are as follows:

- Ingest
- Archival Storage
- Data Management
- Preservation Planning
- Access
- Administration

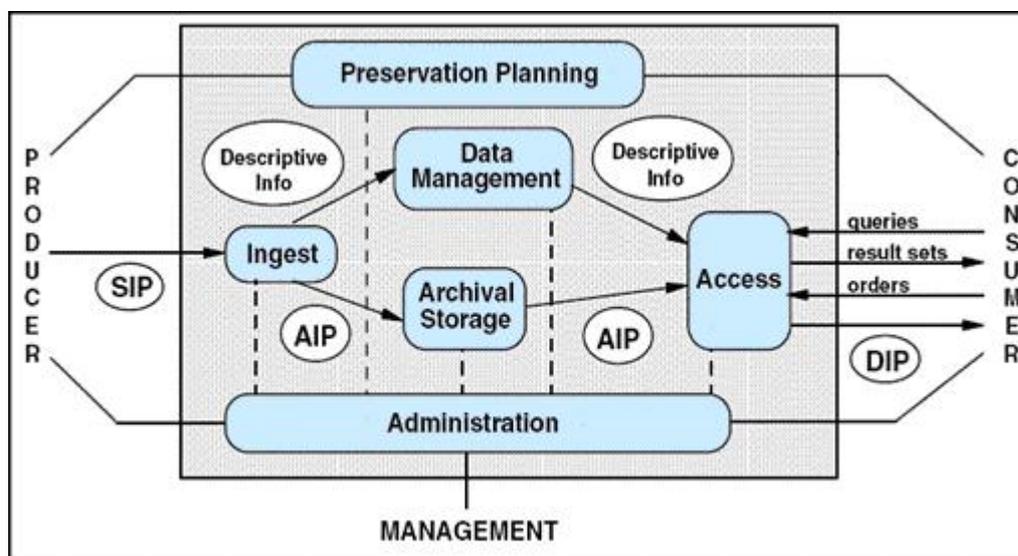


Diagram sourced from the Reference model for an Open Archival Information System (OAIS) CCSDS 650.0-B-1 Blue Book January 2002

These entities will be the basis of today's workshop and each one will be considered through detailed discussion.

## **Some Definitions**

*Information Package:* The Content Information and associated Preservation Description Information, which is needed to aid in the preservation of the Content Information.

*Content Information:* the set of information that is the focus of preservation along with the associated metadata to ensure it remains renderable and understandable to the Designated Community. It is comprised of its Content Data Object and Representation Information, eg a document and its metadata.

*Submission Information Package (SIP):* An information package that is delivered by the Producer to the OAIS for use in the construction of one or more AIPs.

*Preservation Description Information (PDI):* the additional metadata that is necessary for the adequate preservation of the Content Information, which comprises of Provenance, Reference, Fixity and Context Information.

*Archival Information Package (AIP):* is the information package that is stored and preserved within an OAIS. The AIP is made up of information that is preserved and accompanied by a complete set of metadata, which supports the archival systems preservation and access services.

*Dissemination Information Package (DIP):* The Information package, derived from one or more AIPs, received by the Consumer in response to a request to the archive.

Reference: Technology Watch Report, *The Open Archival Information System Reference Model: Introductory Guide*, January 2004.

## **Adoption of the OAIS Model**

ADRI, in its investigation and development of its *Model Plan for an Archival Authority Implementing Digital Recordkeeping and Archiving* (March 2007) adopts the OASI model and makes the assumption that the Reference Model for an Open Archival Information System (OAIS) is used as a basis for an authority's preservation strategy.

The use of OASI as a basis for this workshop continues from the acceptance of the model by ADRI. However it is important to note that it is not the only model available for the preservation of digital records in a virtual archive.

## Session 8:

### Ingest

#### An outline

The Ingest entity provides the services and functions to accept an object(s) (SIP) from a producer and prepare the contents for storage and management within the archive.

In English this means government or client organisations provide digital objects with their associated metadata and these are prepared for storage within the digital archive.

#### What Ingest means in practice

The Ingest entity includes the functions:

- **Receive submission:** The function of providing the appropriate storage capability or devices to receive a SIP.
- **Quality assurance:** The function of validating the successful transfer of the SIP to the staging area
- **Generate archival information package:** The function of transforming one or more SIPs into one or more AIPs
- **Generate descriptive information:** The function of extracting descriptive information from the AIPs and other sources
- **Co-ordinate updates:** The function of transferring AIPs to archival storage and the descriptive information to Data Management

Key points:

- Information created by an organisation is transferred to the archive
- This transfer includes the document/file plus potentially some metadata
- The archive undertakes a quality check to ensure records are appropriate for preservation
- Does the generation of AIP and DIP differ from current practices?

The attached model depicts the ingest stage within the overall framework of the OAIS.

#### Case Study – Public Record Office of Victoria

<http://www.dlib.org/dlib/november07/waugh/11waugh.html>

The Public Record Office Victoria (PROV) commissioned a digital archive in late 2005. During the design and implementation of this digital archive, considerable attention was paid to the ingest function that accessions digital objects into the archive. In particular, the archive was designed to process large transfers, and particular care was taken to support archivists in managing the transfer and handling the inevitable errors.

The article describes the design of the ingest function, and the lessons PROV have learnt about ingest.

**Discuss article, with particular focus on the ingest function, transfer process and bulk processing. What issues are highlighted in article?**

### **Study Resources**

<http://www.prov.vic.gov.au/vers/digitalarchive/development.htm>

<http://www.naa.gov.au/records-management/secure-and-store/e-preservation/at-NAA/index.aspx>

<http://www.prov.vic.gov.au/vers/digitalarchive/who.htm>

<http://www.oclc.org/digitalarchive/>

### **Discussion**

Break into groups to discuss the following points in relation to the ingest function of a virtual archive (one group per point):

- The technology needed for ingest to occur
- The policies and business rules that need to be in place
- The external products required by agencies
- The people and skills that are required

### **Activity**

Discuss how the above relate to how “ingest” is undertaken in archives traditionally. Compare the above elements with business as it occurs in non-digital archives presently.

Using a sample of elements from one particular field, prioritise them based on the following criteria from High to Low (High being where there is no existing rules/procedures etc, Medium requiring integration and modification, and Low being where there is existing rules/procedures etc.

## Session 9:

### Archival Storage

#### An outline

This entity provides the services and functions for the storage, maintenance and retrieval of Archival Information Packages (AIPs).

Archival Storage functions include receiving AIPs from Ingest and adding them to permanent storage, managing the storage hierarchy, refreshing the media on which archive holdings are stored, performing routine and special error checking, providing disaster recovery capabilities, and providing AIPs to Access to fulfill orders. (ISO 14721, 2003:41-42)

#### What Archival Storage means in practice

The Archival Storage entity includes the following functions:

- **Receive Data:** The function of moving the AIP to permanent storage within the archive
- **Manage storage hierarchy:** The function of positioning the contents of the AIPs on the appropriate media
- **Replace media:** Ensuring the capability to reproduce AIPs over time
- **Error checking:** The function of providing statistically acceptable assurance that no components of the AIP are corrupted during any internal Archival Storage data transfer
- **Disaster recovery:** The function of duplicating the digital contents of the archive collection and storing the duplicate in a physically separate facility
- **Provide data:** The function of providing copies of stored AIPs to Access.

More detailed information on each of the above can be located at:

*CCSDS Reference Model for an Open Archival Information System*

<http://public.ccsds.org/publications/archive/650x0b1.pdf>

#### Study Resources

<http://www.prov.vic.gov.au/vers/digitalarchive/development.htm>

<http://www.naa.gov.au/records-management/secure-and-store/e-preservation/at-NAA/index.aspx>

<http://www.prov.vic.gov.au/vers/digitalarchive/who.htm>

<http://www.oclc.org/digitalarchive/>

## **Discussion**

Break into groups to discuss the following points in relation to the archival storage function of a virtual archive (one group per point):

- The technology needed for archival storage to occur
- The policies and business rules that need to be in place
- The external products required by agencies
- The people and skills that are required

## **Activity**

Discuss how the above relate to how archival storage is undertaken in archives traditionally. Compare the above elements with business as it occurs in non-digital archives presently.

Using a sample of elements from one particular field, prioritise them based on the following criteria from High to Low (High being where there is no existing rules/procedures etc, Medium requiring integration and modification, and Low being where there is existing rules/procedures etc.

## Session 10:

### Data Management

#### An outline

This entity provides the services and functions for populating, maintaining, and accessing both Descriptive Information, which identifies and documents archive holdings and administrative data used to manage the archive.

Data Management functions include administering the archive database functions (maintaining schema and view definitions, and referential integrity), performing database updates (loading new descriptive information or archive administrative data), performing queries on the data

#### What data management means in practice

The data management entity includes the functions:

- **Administer database:** maintaining the integrity of the Data Management database
- **Perform queries:** receiving a query request from Access and executing the query to generate a result set that is transmitted to the requester
- **Generate report:** receiving a report request from Ingest, Access or Administration and executing any queries or other processes necessary to generate the report
- **Receive database updates:** adding, modifying or deleting information in the Data Management persistent storage

More detailed information on each of the above can be located at:

*CCSDS Reference Model for an Open Archival Information System*

<http://public.ccsds.org/publications/archive/650x0b1.pdf>

#### Study Resources

<http://www.prov.vic.gov.au/vers/digitalarchive/development.htm>

<http://www.naa.gov.au/records-management/secure-and-store/e-preservation/at-NAA/index.aspx>

<http://www.prov.vic.gov.au/vers/digitalarchive/who.htm>

<http://www.oclc.org/digitalarchive/>

## **Discussion**

Break into groups to discuss the following points in relation to the data management function of a virtual archive (one group per point):

- The technology needed for data management to occur
- The policies and business rules that need to be in place
- The external products required by agencies
- The people and skills that are required

## **Activity**

Discuss how the above relate to how data management is undertaken in archives traditionally. Compare the above elements with business as it occurs in non-digital archives presently.

Using a sample of elements from one particular field, prioritise them based on the following criteria from High to Low (High being where there is no existing rules/procedures etc, Medium requiring integration and modification, and Low being where there is existing rules/procedures etc.

## Session 11:

### Preservation Planning

#### An outline

This entity provides the services and functions for monitoring the environment of the OAIS and providing recommendations to ensure that the information stored in the OAIS remains accessible to the Designated User Community over the long term, even if the original computing environment becomes obsolete.

Preservation Planning functions include evaluating the contents of the archive and periodically recommending archival information updates to migrate current archive holdings, developing recommendations for archive standards and policies, and monitoring changes in the technology environment and in the Designated Community's service requirements and Knowledge Base. Preservation Planning also designs IP templates and provides design assistance and review to specialize these templates into SIPs and AIPs for specific submissions. . Preservation Planning also develops detailed Migration plans, software prototypes and test plans to enable implementation of Administration migration goals.

#### What preservation planning means in practice

The preservation planning entity includes the functions:

- **Monitor designated community:** The function of interacting with archive consumers and producers to track changes in their service requirements and available product technologies
- **Monitor technology:** The function of tracking emerging digital technologies, information standards and computing platforms
- **Develop preservation strategies and standards:** The function of developing and recommending strategies and standards to enable the archive to better anticipate future changes in the designated community service requirements or technology trends
- **Develop packaging designs and migration plans:** The function of developing new IP designs and detailed migration plans and prototypes, to implement Administration policies and directives.

More detailed information on each of the above can be located at:

*CCSDS Reference Model for an Open Archival Information System*  
<http://public.ccsds.org/publications/archive/650x0b1.pdf>

## **Study Resources**

<http://www.prov.vic.gov.au/vers/digitalarchive/development.htm>

<http://www.naa.gov.au/records-management/secure-and-store/e-preservation/at-NAA/index.aspx>

<http://www.prov.vic.gov.au/vers/digitalarchive/who.htm>

<http://www.oclc.org/digitalarchive/>

## **Discussion**

Break into groups to discuss the following points in relation to the preservation planning function of a virtual archive (one group per point):

- The technology needed for preservation planning to occur
- The policies and business rules that need to be in place
- The external products required by agencies
- The people and skills that are required

## **Activity**

Discuss how the above relate to how preservation planning is undertaken in archives traditionally. Compare the above elements with business as it occurs in non-digital archives presently.

Using a sample of elements from one particular field, prioritise them based on the following criteria from High to Low (High being where there is no existing rules/procedures etc, Medium requiring integration and modification, and Low being where there is existing rules/procedures etc).

## Session 12:

### Access

#### An outline

This entity provides the services and functions that support Consumers in determining the existence, description, location and availability of information stored in the OAIS, and allowing Consumers to request and receive information products.

Access functions include communicating with Consumers to receive requests, applying controls to limit access to specially protected information, coordinating the execution of requests to successful completion, generating responses (Dissemination Information Packages, result sets and reports) and delivering the responses to the consumers.

#### What access means in practice

The access entity includes the functions:

- **Coordinate access activities:** Provides a single user interface to the information holdings of the archive
- **Generate DIP:** The function of accepting a dissemination request, retrieving the AIP from Archival Storage and moving a copy of the data to a staging area for further processing
- **Deliver response:** Handles both on-line and off-line deliveries of responses (eg DIPs, reports, assistance) to consumers.

More detailed information on each of the above can be located at:

*CCSDS Reference Model for an Open Archival Information System*  
<http://public.ccsds.org/publications/archive/650x0b1.pdf>

#### Case Study

<http://www.ed.gov/Technology/Futures/honey.html>

The phenomenal growth of the World Wide Web and the rapid development of multimedia technologies are making a wide range of information resources available to the public in ways that were heretofore impossible. Yahoo, the company that provides indexing and search tools for Internet resources, claims that there are 50,000 new Web servers coming online each week, and O'Reilly and Associates, a technology company best known for its Internet-based publications, predicts that the number of adults with Internet access will nearly triple in the next year from 5.8 million to 15.7 million.

New and emerging technologies can be powerful allies in efforts to improve teaching and learning. For technologies to play a truly comprehensive role in

reshaping the nature of schooling, however, they must support a broader vision of educational change. [The Center for Children and Technology](#) has long been committed to a vision of schools and classrooms in which reasoning and problem solving are emphasized, students and teachers help each other learn, students and teachers are motivated and genuinely engaged, the learning environment is responsive to students, teachers use authentic assessment practices, and students learn and effectively apply complex thinking skills and discipline-based knowledge.

This paper is an attempt to begin a conversation about how to make large digital archives genuinely valuable to the K-12 community. The three design considerations addressed here are not meant to be exhaustive, but rather are meant to initiate conversation about how to design resources that are responsive to the needs of elementary and secondary educators. Successful efforts will provide enough structure to make archival collections useful to teachers and students, while preserving the atmosphere of serendipity that is key to imaginative discovery.

### **Study Resources**

<http://www.prov.vic.gov.au/vers/digitalarchive/development.htm>

<http://www.naa.gov.au/records-management/secure-and-store/e-preservation/at-NAA/index.aspx>

<http://www.prov.vic.gov.au/vers/digitalarchive/who.htm>

<http://www.oclc.org/digitalarchive/>

### **Discussion**

Break into groups to discuss the following points in relation to the access function of a virtual archive (one group per point):

- The technology needed for access to occur
- The policies and business rules that need to be in place
- The external products required by agencies
- The people and skills that are required

### **Activity**

Discuss how the above relate to how access is undertaken in archives traditionally. Compare the above elements with business as it occurs in non-digital archives presently.

Using a sample of elements from one particular field, prioritise them based on the following criteria from High to Low (High being where there is no existing rules/procedures etc, Medium requiring integration and modification, and Low being where there is existing rules/procedures etc.

## Session 13:

### Administration

#### An outline

This entity provides the services and functions for the overall operation of the archive system.

Administration functions include soliciting and negotiating submission agreements with Producers, auditing submissions to ensure that they meet archive standards, and maintaining configuration management of system hardware and software. It also provides system engineering functions to monitor and improve archive operations, and to inventory, report on, and migrate/update the contents of the archive. It is also responsible for establishing and maintaining archive standards and policies, providing customer support, and activating stored requests.

#### What administration means in practice

The administration entity includes the functions:

- **Negotiate submission agreement:** The function of soliciting desirable archival information for the archive and negotiating submission agreements with producers. Also the function of negotiating the data submission schedule
- **Manage system configuration:** The function of system engineering for the archive system to continuously monitor the functionality of the entire archive system and systematically control changes to the configuration
- **Archival information update:** A mechanism for updating the contents of the archive
- **Physical access control:** Mechanisms to restrict or allow physical access to elements of the archive, as determined by archive policies
- **Establish standards and policies:** The function of establishing and maintaining archives system standards and policies
- **Audit submission:** The function of verifying that submissions (AIP or SIP) meet the specifications of the Submission Agreement
- **Activate requests:** The function of maintaining a record of event-driven requests and periodically comparing it to the contents of the archive to determine is all needed data available
- **Customer service:** The function of creating, maintaining and deleting customer accounts.

More detailed information on each of the above can be located at:

*CCSDS Reference Model for an Open Archival Information System*  
<http://public.ccsds.org/publications/archive/650x0b1.pdf>

## **Study Resources**

<http://www.prov.vic.gov.au/vers/digitalarchive/development.htm>

<http://www.naa.gov.au/records-management/secure-and-store/e-preservation/at-NAA/index.aspx>

<http://www.prov.vic.gov.au/vers/digitalarchive/who.htm>

<http://www.oclc.org/digitalarchive/>

## **Discussion**

Break into groups to discuss the following points in relation to the administration function of a virtual archive (one group per point):

- The technology needed for administration to occur
- The policies and business rules that need to be in place
- The external products required by agencies
- The people and skills that are required

## **Activity**

Discuss how the above relate to how administration is undertaken in archives traditionally. Compare the above elements with business as it occurs in non-digital archives presently.

Using a sample of elements from one particular field, prioritise them based on the following criteria from High to Low (High being where there is no existing rules/procedures etc, Medium requiring integration and modification, and Low being where there is existing rules/procedures etc).

## Session 14:

### Conclusion

During the past two days, a broad variety of concepts, issues and problems have been addressed. These have ranged from the traditional concepts of recordkeeping and archiving, through to current, state of the art case studies of digital archiving and related models.

To finalise our workshop today, take the time individually to think about what you consider to be the leading issues/problems/challenges that lie ahead for your respective archival institution.

Of these issues/problems/challenges, which of them can we change/resolve as professionals in our field?

What is it about those problems/issues/challenges that have been left behind, that we somehow feel that we can't change?

(Responses will include "other people", "out of our hands", "beyond our means and resources" etc)

Lets look at these with the time we have left. If we feel we cannot change them, perhaps we should be thinking about how we can influence changing them. With that frame of mind, lets reconsider them...how many of these can we "influence"?