ADRI

Digital Record Export Standard

ADRI Submission Information Package (ASIP)

ADRI-2007-01-v1.0

Version 1.0
31 July 2007

Version	Date	Author	Comment
1.0	31.07.2007		First released version

Endorsement

This document has been endorsed by the Australian Digital Recordkeeping Initiative as a standard on 4 October 2006.

This document has been issued by the Council of Australasian Archives and Records Authorities as a standard on 31 July 2007.

Acknowledgements

We would like to acknowledge the help given by those people who commented on drafts of this document.

Table of Contents

1 Overview of the Standard

This standard describes requirements for a transfer package of digital records to an archive. It specifies:

- the structure of a digital transfer package,
- the metadata that must be transferred with each record
- optional metadata that may be transferred with each record

The purpose and membership of ADRI as well as its relationship to its parent body CAARA is described in section 1.1. The ADRI Framework for digital recordkeeping, in which this standard fits, is presented in section 1.2. The background to the standard and more detail as to what this standard covers is given in section 1.3 and 1.4.

The standard does not cover the:

- high level transfer process that negotiates and manages the transfer behavior of the records system[1]
- metadata requirements of a records system beyond that required for export
- format of the record content
- low level protocols that ensure complete and accurate transfer
- physical transfer mechanisms

These exclusions are described in more detail in section 1.5.

We expect the standard will be of use to those:

- responsible for records within an organization
- carrying out transfers of digital records
- designing or implementing records systems that must conform with this standard
- testing records systems for conformance to this standard.

These roles are described in more detail in section 1.6.

The nature of the standard means that there may be differing implementations of the standard. This is described in section 1.7.

ADRI is beginning a collaborative, multi year, effort to develop a revised Digital Record Export Standard that will take into account jurisdictional requirements and international developments. The results of this project will be issued as Version 2 of this standard.

---

[1] In this standard we will distinguish between a *recordkeeping system* (an overarching structure encompassing myriad individual systems – both business and records specific – people, and resources), *records systems* (those applications which make and/or manage records). Records systems are further broken down into *business systems* that have a records creation and/or management component, and *records application software* which is dedicated to managing records.

1.1   The Australasian Digital Recordkeeping Initiative (ADRI)

The Australasian Digital Recordkeeping Initiative (ADRI) is composed of representatives from all state and national archival authorities in Australia and New Zealand. The members of ADRI are:

- National Archives of Australia
- Archives New Zealand
- Public Record Office Victoria
- State Records NSW
- ACT Territory Records
- Archives Office of Tasmania
- Northern Territory Archives Service
- Queensland State Archives
- State Records South Australia
- State Records Office Western Australia

The aim of the Initiative is to develop and harmonise a uniform set of standards, guidelines, and practices for digital recordkeeping. A related aim is to improve the organisational capability, capacity, and expertise within the collaborating institutions in relation to digital recordkeeping.

ADRI is a working group of the Council of Australasian Archives and Records Authorities (CAARA). CAARA is the peak body of government archives and records institutions in Australia and New Zealand.

1.2   ADRI Framework – a uniform approach to electronic recordkeeping

The members of ADRI have agreed to collaborate on the development, articulation, and implementation of a common set of strategies for enabling the making, keeping and using of digital records. ADRI will promote a single Australasian approach to digital public recordkeeping across all jurisdictions and provide a space for communication and information sharing between the members. The collaboration builds on and acknowledges many years of Australasian collaboration in the development of concepts, tools, standards, and strategies for good recordkeeping.

The collaborative approach to digital recordkeeping adopted by ADRI consists of a number of inter-related component parts. ADRI has prepared a Framework document[2] to provide a high-level overview of these component parts of the Australasian approach to digital recordkeeping in the 21st Century. Many of these components will have jurisdiction-specific manifestations that, while reflecting local circumstances and requirements, will all nevertheless conform to the fundamental features of the common framework. This Standard relates to that part of the Framework that is

---

[2] ADRI Framework 2005-2007, available from http://www.adri.gov.au/subcontent.asp?scID=7 visited 1 June 2007.

specifically concerned with the transfer of digital objects between records systems and between producers and managers of archival records.

The ADRI Framework consists of four broad components:
- making and managing digital records.
- keeping digital records in agencies and in archives and records authorities
- transferring digital records to archives and records authorities
- using digital records

All of the components contain guidelines and standards. This standard forms part of the standards associated with the transfer component. The other standard associated with the transfer component deals with minimum authenticity requirements. The guidelines associated with the transfer components are:
- Preferred data formats and methods for transferring records to archival custody
- Methods for automatic transfer of recordkeeping metadata
- Maintenance of provenance and authenticity

1.3   Background to this Standard

The purpose of this standard is to simplify and reduce the cost of transferring digital records from agencies to member archives. This is achieved by specifying a standard representation for a digital record. The records are expressed in this standard format, known as a Submission Information Package (SIP), by agencies before transfer.

Records are evidence of an organisation or individual carrying out their day-to-day activities. Understanding what happened in the past and why is critical to the continued day-to-day activities of an organisation or individual. It is also the basis on which our legal system is built, and it is the basis of all historical understanding. For all these reasons, preserving the ability to access records is critical to organisations and individuals.

Electronic records are simply those represented electronically; today this almost exclusively means represented digitally. Electronic records range from written documents (such as reports and emails), to images (photographs, plans), sound (voice mail), video, and databases. The continued integration of the computer into infrastructure means that new types of electronic records are being continually developed.

An electronic recordkeeping encompasses recordkeeping metadata relating to content, context, and structure. Metadata establishes the record's authenticity, integrity, and reliability.

In an organisation electronic records are stored and managed by many different records systems. These systems may range from the very simple (e.g. the corporate shared file servers) through the sophisticated (e.g. a dedicated electronic document management system which provides

sophisticated record management functionality), to the extremely complex (e.g. custom built business applications that store specialised records). All of these records systems have one characteristic in common: they have a finite, relatively short life. Often a records system has a much shorter life than the records that it holds.

The short lifespan of records systems leads to a key challenge in preserving electronic records: ensuring that the electronic records can be extracted from the records systems that currently store and manage them and be reliably transferred to another system. Three transfer instances are of particular importance:

- Transfer of records from one system to its replacement within an organisation. This type of transfer occurs routinely, albeit at a relatively low frequency.
- Transfer of long term temporary records from the records system that created them to a specialist storage system (which may be operated by a separate organisation) and back again. This is equivalent to the practice of secondary storage of paper records. This relieves an organisation from having to preserve and store records that have no day-to-day use in operational systems.
- Transfer of permanent records from the records system that created them to a specialist archival records system (which may be operated by a separate organisation). This ensures that long-term preservation of permanent records is carried out in a specialist system and relieves the creating system from the burden of preserving, storing, and providing access to records.

A Submission Information Package (SIP) can be used as a lingua franca between records systems. Without such a standard, any transfer of electronic records between records systems requires specific custom arrangements between each pair of systems. Such custom arrangements have a:

- high risk of loss or compromised records
- high cost

The high risk and cost can often be managed in the limited case where a records system is superseded within an organisation by another system. In general, however, managing the risk and cost is far more difficult where an organisation (such as an archive or secondary storage supplier) is receiving records from many organisations.

A SIP standard allows the risk and cost of transfer to be managed.

The use of a SIP standard means that less custom software is required to transfer records, and this software can be re-used. Re-use may occur by the same organisation (e.g. a secondary storage supplier using the same software to accept transfers from many different organisations), or by different organisations (e.g. a vendor of a recordkeeping system implementing the transfer software once and providing it to all users of the system).

The use of a SIP standard also means that less intervention is required during transfer. Intervention may occur at export, physical transfer, and import of records.

In other words a standardised SIP is part of risk minimisation and therefore ultimately results in cost reduction.

1.4 Goal of the Standard

The goal of this standard is to define a standard SIP that may be used by all records systems in all organisations when transferring records between systems. The systems may be located in the same organisation or in different organisations. In particular, the SIP can be used when exporting records to an archive.

The SIP standard is designed to reduce the risk of loss or compromised records, and the overall cost of the transfer. A SIP standard supports this transfer by defining a standard representation for an electronic record.

1.5 Features Out of Scope of the Standard

This standard will not define:
- The high level transfer process between two records systems or between an agency and an archive. This process involves, among other things, the negotiation about what is to be transferred and when (possibly between two organisations); the monitoring of the transfer to ensure that quality and cost parameters are met; and the checking that the transfer was carried out correctly.
- The behaviour of a records system while holding electronic records. The standard assumes that the records system is designed to hold records in a manner that ensures their authenticity and integrity.
- Recordkeeping metadata held by a records system that is not transferred to an archive.
- The data formats (e.g. PDF, TIFF) in which an archive will accept the content of electronic records. The SIP is designed to contain any content of any format. An archive may, outside this standard, restrict the data formats that it accepts.
- The low level protocols that ensure that all records are correctly transferred and custody accepted by the archive before an agency disposes of its copy of the record.
- The physical mechanisms by which the records are transferred from one records system to another or from an agency to an archive. This standard defines the physical representation of a record outside a records system. This physical representation could be moved between the records systems in many different ways (e.g. by the internet or on media such as CDs). The physical transfer mechanisms will not be defined in this standard.

1.6   Audience for the Standard

The audience for this standard is:
- Those responsible for records within organisations (e.g. recordkeepers, archivists, and their managers). This group needs to understand the purpose of the standard and, broadly, how it is achieved. This group also need to know what the standard does not cover so that these areas can be negotiated prior to a transfer.
- Those responsible for carrying out a transfer of records between systems (e.g. recordkeepers, archivists). This standard enables them to know how a record is represented so that they can carry out a specific transfer.
- Those responsible for designing and implementing software for exporting and importing records from/to records systems (e.g. vendors, software developers). This standard enables them to have a precise, clear, and unambiguous specification of the SIP so that they can accurately implement it.
- Those responsible for testing systems to ensure that systems accurately implement the SIP. This standard provides a precise, clear, and unambiguous specification, and test tools to ensure defensible decisions about whether an implementation fulfils the standard or not.

1.7   Implementing this Standard

It is envisaged that this standard may be implemented in a number of ways. An example of one way of implementing is included in Appendix I but there may in future be other compliant implementations (e.g. using a jurisdictional specific metadata standard different to that outlined in Appendix I).

While all ADRI members have endorsed this standard, not all members of its parent body CAARA will accept every implementation of the standard (including the implementation given in Appendix I) for transfer into their own archival custody. Implementers should check with the individual CAARA member about which implementations they will accept before undertaking any implementation work for use in a specific CAARA member's jurisdiction.

2 References

2.1 OAIS Reference model (ISO 14721:2003)

The Open Archival Information System reference model[3] provides a conceptual framework for the provision of a persistent archive of digital information. A persistent archive ensures that access can be provided to the contents of the archive in a reliable and consistent way.

This standard references the model in so far as it assumes that the components of the model will be available to support the management of the transfer package format specified here.

A key component of the OAIS reference model is the Submission Information Package (SIP). A SIP contains the digital information and associated metadata delivered from the producer of the information to the OAIS (archive) during the transfer.

This standard specifies the requirements for a transfer package format as the standard implementation of the OAIS Submission Information Package (SIP) for the ADRI jurisdictions.

---

[3] Space data and information transfer systems – Open archival information system – Reference model, ISO 14721:2003. This is also available as 'Reference Model for an Open Archival Information System (OAIS)', Blue Book, Issue 1, January 2002, CCSDS 650.0-B-1, Consultative Committee for Space Data Systems.

3   ADRI SIP Requirements

This section defines requirements that all ADRI Submission Information Packages (ASIPs) must support.

3.1   General

The specification of a package accepted as an ASIP must:
* Be freely available to any interested party. The freely available copy may be made available electronically (e.g. via the Web) or via a printed copy. 'Freely available' does not necessarily mean free; a reasonable charge may be made for the specification.
* Have a formal compliance process so that implementations of the specification can be determined to be compliant or otherwise.

3.2   The Submission Information Package (SIP)

For the purposes of this standard, the basic component of an ADRI SIP (ASIP) is a digitally encoded package of one or more records. Each ASIP represents one or more digital records. A transfer between a creator and an archive may contain one or more separate digital records, and hence comprise one or more ASIPs.

The structure of a record may vary between archival institutions. For some institutions a record may consist of a single document with one encoding. Other institutions have a more complex structure; for example, a document may contain multiple encodings, or a record may consist of multiple documents.

An ASIP consists of:
* A declaration that this is an ADRI SIP (see section 3.3.1)
* An authentication method to detect tampering to the body of the ASIP (see section 3.5)
* Metadata about the record (see sections 3.3.2 and 3.3.3)
* The actual content of the record

3.3   Metadata

The purpose of this section is to define the metadata that may occur within an ADRI SIP.

The metadata defined in this standard are grouped according to the following categories:
* Packaging Information (high level descriptive information such as the declaration, ASIP Type etc);
* Preservation Description Information (information that allows the content information to be understood, including unique identifiers, contextual and provenance information, and fixity information);

- Content Information (the actual bits to be preserved and how they are to be interpreted);

A recordkeeping system must be capable of generating the mandatory metadata defined in this standard when exporting an ASIP.

3.3.1 Packaging Information metadata

This records ASIP-level metadata that must be included to identify, delimit, and describe an ASIP as a whole. Note that this information does not describe the records held within the ASIP.

The metadata captured at this level may include:
- ASIP Declaration
- ASIP Format Description
- ASIP Type Description
- ASIP Creation Date
- ASIP Content

3.3.2 Preservation Description Information metadata

This contains information about the record contained within the ASIP. The information includes:
- Reference information. This information contains identifier(s) of the record. Typical reference information in an archival environment includes an agency identifier, a series identifier, and a record identifier.
- Context information. This information contains the context of the record; that is, information describing this record and its relationship to other records and its environment.
- Provenance information. This information describes the history of the record.
- Fixity information. This information is used to ensure that the record content has not been modified in an undocumented fashion. Typical fixity information includes: algorithm identifiers, checksums or signatures, and support information such as certificates.

The ASIP must specify which recordkeeping metadata standard is being used for contextual and provenance information, e.g. the ASIP may declare the NAA Recordkeeping Metadata Standard, or may declare the New South Wales Recordkeeping Metadata Data Standard. An archival institution may choose to accept only one recordkeeping metadata standard, in which case the specification is implicit.

3.3.3 Content Information metadata

This metadata group contains metadata elements relating to the record content. This group also contains the bits that are being preserved.

The elements captured at this level can include:
- Format encoding

- Source File Identifier
- Rendering Keywords
- Document Data

3.4   Encryption

NO PART of the ASIP can be encrypted, except during the transmission between the producer and the archive. If an ASIP is encrypted during transmission, full decryption must be arranged with the receiving archive.

3.5   Authentication Requirements

All ASIPs must include authentication mechanisms for both the source of the ASIPs (i.e. that the ASIPs have come from the designated Producer) and the integrity of the ASIPs (that the ASIPs have not been corrupted).

Appendix I   Victorian Electronic Record Strategy (VERS)

The following two specifications that form part of PROS 99/007 (PROV Standard *Management of Electronic Records,* commonly known as the 'VERS' standard) jointly fulfil the functionality required by this standard:

- *Management of Electronic Records (PROS 99/007) version 2, Specification 2, VERS Metadata Schema*
- *Management of Electronic Records (PROS 99/007) version 2, Specification 3, VERS Standard Electronic Record Format*

These components are modified by the *Errata for Management of Electronic Records PROS 99/007 (Version 2)*

The two specifications and the errata are available from http://www.prov.vic.gov.au/vers/standard/version2.htm and the compliance tests are available on http://www.prov.vic.gov.au/vers/compliance/

The remaining three specifications that form part of PROS 99/007 are not relevant to ASIP and do not form part of this standard:

- *Specification 1, System Requirements for Preserving Electronic Records*
- *Specification 4, VERS Long Term Preservation Formats*
- *Specification 5, Export of Electronic Records to PROV*