

HDMS Application Guidelines for EAD

Introduction

The Heritage Documentation Management System provides a suite of tools for an archivist to process and manage any collection or group of records. The HDMS has an HTML finding aid generator to produce a multi-page HTML finding aid, which includes links between series, provenance, inventory, related series, and related provenance. An index and photo gallery pages can also be added.

In this context, the aim of the EAD generator for the HDMS is to provide a basic SGML or XML EAD finding aid in the first instance. The SGML finding aid produced may be uploaded into an appropriate SGML viewer, and the XML finding aid can be viewed, via an XSL stylesheet in an XML enabled browser, i.e. Internet Explorer 5*. Thus, not all fields in the HDMS are output to EAD, nor are all the possible mappings of HDMS fields to EAD used. The aim is to strike an appropriate balance, and produce a structured finding aid, that is not 'overloaded' with data, nor is 'too large' to be delivered over the web.

These guidelines detail how fields in the HDMS are used to produce an EAD finding aid. The guidelines are based on guidelines for other EAD projects, e.g. *RLG Recommended Application Guidelines for EAD* (<http://www.rlg.org/flgead/guidelines.html>) and attendance at an EAD workshop at the NLA in Canberra, February 28 to March 1 2000 for the National Database of Electronic Finding Aids for Australian Literary Manuscript Collections REIF Project. Reference has also been made to a basic XML text, namely St. Laurent, Simon, *XML A Primer*, MIS:Press, California, 1998, ISBN 155828592X.

I. General Notes

Nomenclature

- **tag** - markers that enclose the content of an element, e.g. <p>[content]</p>
- **element** - textual concept thereby represented
- **content** - what is enclosed in the tag
- **value** - attribute options

(Source: *RLG Recommended Application guidelines*)

HDMS field names are represented using **tablename![fieldname]**, e.g. **Inventory![Titleins]** refers to the Titleins field of the Inventory table.

Data from the following HDMS tables are used in the production of a finding aid:-

- ASPI (Accession, Series, Provenance and Inventory) tables which document the records,
- ASPI profile tables which hold summary information calculated from the ASPI tables,
- FindingAid which holds information about the finding aid - e.g. boilerplate, content options,
- HTML which holds information specific to creation of an HTML Finding Aid,
- EAD which holds information specific to creation of an EAD Finding Aid, and
- METADATA which holds Dublin Core and other metadata information.

Overview of EAD structure

EAD is made up of three discrete parts:

- **<eadheader>** - information describing the finding aid itself
- **<frontmatter>** - prefatory matter incorporating information useful for the display or publication of the finding aid

* IE5 supports an earlier version of XSL, i.e. <http://www.w3.org/TR/WD-xsl> as opposed to the latest version, <http://www.w3.org/1999/XSL/Transform>

- **<archdesc>** - the description of the archival materials and associated administrative and contextual information

(Source: RLG Recommended Application guidelines)

SGML versus XML

The HDMS EAD generator allows for the creation of SGML EAD documents and/or XML EAD documents. The differences are-

- a) Opening prolog - XML requires XML processing instruction, link to stylesheet and entity declaration for EAD notations, see II below for more details.
- b) uppercase tags for SGML (a stylistic decision).
- c) lower case tags for XML (necessary as XML is case sensitive).

The files making up the SGML EAD DTD are available for download from the official EAD web site at <http://lcweb.loc.gov/ead/eadv1ann.html>. Modifications need to be made to make them suitable for 'XML'.

II DOCTYPE and Entity Declarations

The document type declaration is used to indicate the type of document and to what document type definition (DTD) it adheres. It begins with `<!DOCTYPE` followed by the name of the DTD, followed by a link pointing to where the DTD may be found.

(Source: XML A Primer)

As the EAD DTD is a public standard, the PUBLIC keyword is used followed by the public identifier, in quotes, and then the link to where the DTD may be found.

Entity Declarations are any general or parameter entities, not defined in the DTD, that need to be defined for the particular EAD finding aid being produced. For more information on entities and their uses, see any SGML/XML text.

In the HDMS, the location of the DTD and any entity declarations are recorded in the DTDLocation and EntityDeclarations fields of the EAD table. The location of the DTD should include the path (relative, absolute or URL) to the DTD document when the EAD guide is published on a Web server.

SGML document type declaration

```
<!DOCTYPE EAD PUBLIC "-//Society of American Archivists//DTD ead.dtd (Encoded Archival Description (EAD) Version 1.0//EN" "EAD![DTDLocation]" [ EAD![EntityDeclarations] ]>
```

XML document type declaration

For an XML EAD document, the opening prolog must include the `<?xml?>` processing instruction, a link to the stylesheet, the document type declaration (with lower case ead), and entity declaration for EAD Notation Declarations

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="EAD![XSLFile]"?>
<!DOCTYPE ead PUBLIC "-//Society of American Archivists//DTD ead.dtd (Encoded Archival Description (EAD) Version 1.0)//EN" "EAD![XMLDTDLocation]" [ <!ENTITY % eadnotat PUBLIC "-//Society of American Archivists//DTD eadnotat.ent (EAD Notation Declarations)//EN" "eadnotat.ent"> EAD![EntityDeclarations] ]>
```

Metadata

In the HDMS EAD finding aid, Dublin Core metadata elements are added as a comment after the doctype and entity declarations. The metadata elements are designed to synchronize with the appropriate fields in the FindingAid table, i.e. you change information in the FindingAid title field and the appropriate DC metadata field is updated. Our mapping of elements is in Appendix B. Note that these mappings are based on providing Dublin Core metadata for the finding aid as a published digital object, and not mixing

metadata that applies to the records alone. For example the DC.creator tag does not include the records creator/originator/provenance, as while they may be responsible for creating records, they have not played this role in creating the finding aid.

Mapping of Dublin Core to EAD is provided as an appendix in the EAD Application Guidelines, available in print from the SAA, but unfortunately not available on the web.

III EAD HEADER <eadheader>

The <eadheader> element contains information describing the EAD finding aid. It is a required element, but typically content is not displayed to users. It can be used for retrieval, e.g. contents can be used in providing a brief citation to the finding aid in a search result set, and also to control the EAD instance, i.e. setting of audience and findaidstatus attributes.

```
<eadheader audience="EAD![Audience]" langencoding="iso 639-2"
findaidstatus="EAD![FindAidStatus]">
  <eadid type="SGML/XML catalog">PUBLIC "-// EAD![CountryCode]: : EAD![OrgCode]//TEXT
(EAD![CountryCode]: : EAD![OrgCode]: : FindingAid![RepRef]: : FindingAid![Title]//EN"
"EAD![Filename]/EAD![XMLFilename]"</eadid>
  <filedesc>
    <titlestmt>
      <titleproper>FindingAid![Title]</titleproper>
      <author>Listed by FindingAid![Prepby] for FindingAid![Creator]</author>
      <sponsor>formulated from HTMLSponsors table</sponsor>
    </titlestmt>
    <publicationstmt>
      <publisher> FindingAid![PubBy]</publisher>
      <date> FindingAid![PubDate]</date>
    </publicationstmt>
    <editionstmt>
      <edition>SGML/XML EAD edition</edition>
    </editionstmt>
  </filedesc>
  <profiledesc>
    <creation>Machine readable finding aid generated from the Heritage Documentation
Management System on <date>FindingAid![DateModified]</date></creation>
    <langusage>Finding aid is written in
    <language>FindingAid![Language]</language></langusage>
  </profiledesc>
</eadheader>
```

IV FRONT MATTER <frontmatter>

The <frontmatter> element contains title or front page information. In the HDMS xsl stylesheet it will be displayed as a title page section at the top of the file. Note that the same fields that are used in the <eadheader> element are used in the same manner in the <frontmatter> element. This is because it is the same 'element of information', just used in two different places. The beauty of the database is that you only need to input it once, and it can be reproduced wherever required.

```
<frontmatter>
  <titlepage>
    <num> FindingAid![RepRef] </num>
    <titleproper> FindingAid![Title] </titleproper>
    <publisher> FindingAid![PubBy] </publisher>
    <author> Listed by FindingAid![Prepby] for FindingAid![Creator] </author>
    <sponsor>formulated from HTMLSponsors table</sponsor>
    <date> FindingAid![PubDate] </date>
    <edition>SGML/XML EAD edition</edition>
    <p> FindingAid![Copyright] </p>
  </titlepage>
</frontmatter>
```

V. ARCHIVAL DESCRIPTION <archdesc>

The <archdesc> element contains description of the archival materials and associated administrative and contextual information. It consists of four components:-

- Descriptive Identification <did> - 'the high level <did>' to describe the collection as a whole, critical for meaningful searching and consistent display
- Descriptive Context Elements, e.g. <scopecontent>, <bioghist>, <controlaccess>, <admininfo>, <odd>
- Description of Subordinate Components <dsc> - series descriptions and container lists
- Adjunct to Descriptive data <add>

a) DID

```
<ARCHDESC LEVEL=" FindingAid![ArchDescLevel]" LANGMATERIAL=" FindingAid![LanguageOfMaterial]" >
  <DID>
    <HEAD>About the records</HEAD>
    <UNITTITLE LABEL="Title">Records of Organisation![OrganizationName]</UNITTITLE>
    <UNITDATE LABEL="Date Range">[Inventory Profile]![Earliest] - [Inventory Profile]![Latest]
    </UNITDATE>
    <UNITID LABEL="Reference Number">FindingAid![RepRef]</UNITID>
    <ORIGINATION LABEL="Provenance"><REF TARGET=" FindingAid![PrimaryProv_id]"><CORPNAME
    or PERSNAME>Provenance![N](Provenance![PSTART]- Provenance![PEND])</CORPNAME or
    PERSNAME></REF></ORIGINATION>
    <PHYSDESC LABEL="Extent">[Inventory Profile]![Linear]/100 metres ([Inventory Profile]![qn]
    items)</PHYSDESC>
    <REPOSITORY LABEL="Repository"> FindingAid![Repository] </REPOSITORY>
    <ABSTRACT LABEL="Abstract"> FindingAid![Abstract] .</ABSTRACT>
  </DID>
  <ADMININFO> ... </ADMININFO>
  <SCOPECONTENT> ... <SCOPECONTENT>
  <BIOGHIST> ... </BIOGHIST>
  <DSC> ... </DSC>
</ARCHDESC>
```

b) Descriptive Contextual Elements

- Administrative Information* - formulated using same structure as About the records in the HDMS HTML Finding Aid.

```
<ADMININFO>
  <HEAD>Administrative Information</HEAD>
  <ACQINFO>
    <HEAD>Accession Information</HEAD>
    <P>The collection is comprised of records from [Accession Profile]![Accession no]
    accessions. The codes used to uniquely identify each accession range from
    [Accession Profile]![Accession min] to [Accession Profile]![Accession max] .</P>
  </ACQINFO>
  <PROCESSINFO>
    <HEAD>Processing Information</HEAD>
    <P>The records have been allocated to [Series Profile]![Series no] series. The codes
    used to uniquely identify each series range from [Series Profile]![Series min] to
    [Series Profile]![Series max] .</P>
    <P>Through the processing of the records, [Provenance Profile]![Provenance no]
    provenance entities were identified.
    The codes used to uniquely identify each proveniential entity, i.e. records
    creator or custodian, range from [Provenance Profile]![Provenance min] to [Provenance
    Profile]![Provenance max] .</P>
    <P>The inventory covers [Inventory Profile]![qn] items, and may include records of
    continuing value, records sentenced for destruction and records that have been
    destroyed. The codes used to uniquely identify each inventory item range from
    [Inventory Profile]![Imin] to [Inventory Profile]![Imax] . The total collection occupies
    [Inventory Profile]![linear] linear cm of shelf space (or its equivalent).</P>
    <P>The documentation of the records at inventory level started on [Inventory
    Profile]![Istart] . The latest additions were made on [Inventory Profile]![Ipad] . The
```

```

    latest modifications were made on [Inventory Profile]![Ipmodd]. This collection
    profile was updated on [Inventory Profile]![update].</P>
  </PROCESSINFO>
</ADMININFO>

```

- *Scope and Content* - added to HTML Finding Aid

```

<SCOPECONTENT>
  <HEAD>Scope and Content</HEAD>
  <P><b>Organisation![aboutorg]</b></P>
</SCOPECONTENT>

```

- *Biographical History* - output of all the provenance entities. Primary provenance is output first with <HEAD> data of Biographical Note if the provenance entity is person or family, History otherwise, with rest of the provenance entities output as nested <BIOGHIST> elements.

```

<BIOGHIST ID="Provenance![Prov_id]">
  <HEAD>Biographical Note or History</HEAD>
  <P><TITLE RENDER="bold">Provenance![N]</TITLE></P>
  <P>Provenance![PsubName]</P>
  <CHRONLIST>
    <CHRONITEM><DATE>Provenance![Pstartdate]</DATE><EVENT>Established/Born</EVENT></CHR
    ONITEM>
    <CHRONITEM><DATE>Provenance![Penddate]</DATE><EVENT>Ceased/Died</EVENT></CHRONITEM
    >
  </CHRONLIST>
  <P>Provenance![PSumNote]</P>
  <TITLE RENDER="bolditalic">Full Note</TITLE>
  <P>Provenance![PFullNote]</P>

  <BIOGHIST ID="Provenance![Prov_id]">
    <P>Provenance![PsubName]</P>
    <CHRONLIST>
      <CHRONITEM><DATE>Provenance![Pstartdate]</DATE><EVENT>Established/Born</EVENT></C
      HRONITEM>
      <CHRONITEM><DATE>Provenance![Penddate]</DATE><EVENT>Ceased/Died</EVENT></CHRONIT
      EM>
    </CHRONLIST>
    <P>Provenance![PSumNote]</P>
    <TITLE RENDER="bolditalic">Full Note</TITLE>
    <P>Provenance![PFullnote]</P>
  </BIOGHIST>
  repeat <BIOGHIST> ... </BIOGHIST> for each non primary Provenance
</BIOGHIST>

```

Note - The output of the full note is optional.

<bioghist> Limitations

From the perspective of the HDMS, <bioghist> gives a very primitive representation of provenance entities. Not having the structure to 'title' and 'date' the provenance entity is of particular concern. We have used generic title fields in order to title each provenance entity and also made assumptions in the XSL style sheet to display the born/died, established/ceased date in a 'unitdate' way.

The ability to express provenance entities in a structured way and also to express relationships amongst provenance entities is critical for Australian practice.

c) Description of Subordinate Components <dsc> - series descriptions and container lists

```

<DSC TYPE="combined">
  <HEAD>Series List and Descriptions</HEAD>
  <C01 LEVEL="series" ID="Series![Series_ID]">
    <DID>
      <UNITTITLE>Series![Stitleins] [Series![Stitlequ]]</UNITTITLE>

```



```

<UNITDATE>Series![Sstartdate] - Series![Senddate]</UNITDATE>
<UNITID>Series![Series_ID]</UNITID>
<ORIGINATION LABEL="Provenance"><REF TARGET="Series![Prov_ID]"><CORPNAME or
PERSNAME> Series![Sprovenance]</CORPNAME or /PERSNAME></REF></ORIGINATION>
<PHYSDESC> Series![Squantity] cm, Series![Squantityn] Series![Squantity]</PHYSDESC>
</DID>
<SCOPECONTENT>
  <P>Series![Ssumnote]</P>
</SCOPECONTENT>
<ARRANGEMENT>
  <P>Series![Sarrange]</P>
</ARRANGEMENT>
if Series![SACCESS] <> "0" then
<ADMININFO><ACCESSRESTRICT>
  <P>Closed/Partially closed until Series![Saccessdat]</P>
</ACCESSRESTRICT></ADMININFO>
<CO2 LEVEL="item"> ... </CO2>
... repeat for each item
</CO1>
... repeat for each series
</DSC>

```

Options in the HDMS HTML finding aid to order series by Id or alphabetically by title, to remove codes (i.e. display as Series 1, rather than Series ARRAS0001) and how to treat series where access is closed, are applicable to the EAD finding aid. These options are set by the fields **OrderSeriesbyId**, **RemoveSPCodes** and **ConfSeries** in the Finding Aid Table.

For the treatment of closed series, the **ConfSeries** field can have three values:-

1. **default** - series information is output in the <c01> element , but no inventory listing, i.e. <c02> elements for that series.
2. **exclude** - no series information for series where access is C are included in the output.
3. **include** - treat closed series the same as other series.

Related Series

The HDMS allows for relationships amongst series to be recorded in a RelatedSeries table. Information captured includes the relationship between the series, e.g. the standard relationships of Australian practice - previous, subsequent, controlling, controlled, related - and the ability to define your own, a description of the relationship and the dates of the relationship.

From this information network and hierarchical representations of series may be built. EAD allows for hierarchy, through recursive <c> elements in the <dsc>, but can it support a network model to the same extent? This issue is discussed in the project report *REEA - Building an Encoded Archival Description generator for the HDMS*.

Inventory item description

```

<C02 LEVEL="item">
  <DID>
    <UNITTITLE>Inventory![Titleins] [Inventory![Titlequl]]</UNITTITLE>
    <UNITDATE>Inventory![Startdate] - Inventory![Enddate]</UNITDATE>
    <UNITID>Inventory![Item_id]</UNITID>
    <ORIGINATION LABEL="Provenance"><REF TARGET="Inventory![Prov_id]"><CORPNAME or
PERSNAME> Inventory![Iprovenanc]</CORPNAME or PERSNAME></REF></ORIGINATION>
    <PHYSDESC>
      <EXTENT> Inventory![quantity] cm, Inventory![quantity] Inventory![quantity]</EXTENT>
      <GENREFORM>sentence formulated from Inventory format fields</ GENREFORM >
    </PHYSDESC>
    <CONTAINER LABEL="unit"> Inventory![Unitnum]</CONTAINER>
    <CONTAINER LABEL="series box/box"> Inventory![Boxnum]</CONTAINER>
  </DID>
  <SCOPECONTENT>
    <P> Inventory![Titledet]</P>

```



```
<P>Previous control: Inventory![Control]</P>
</SCOPECONTENT>
if Inventory![ACCESS] <> "O" then
<ADMININFO><ACCESSRESTRICT>
  <P>Closed/Partially closed until Inventory![Accessdate]</P>
</ACCESSRESTRICT></ADMININFO>
</C02>
```

Options in the HDMS HTML finding aid to order the inventory listing, for the label 'Series Box number' rather than 'Box number', to include the physical location and how to treat inventory items where access is closed, are applicable to the EAD finding aid. These options are set by the fields **OrderInv**, **SeriesBoxes**, **BoxLoc** and **Conflnv** in the Finding Aid Table.

For the ordering of inventory, the **OrderInv** field can have four values:-

1. **unitnum** - orders by unitnum (sequence) field.
2. **item_id** - orders by item id, unitnum is not output.
3. **title** -order alphabetically by titleins, unitnum is not output. Note that titles begin with punctuation will appear at the beginning of listing.
4. **date** -order by earliest to latest dates, unitnum is not output. Note items without dates will appear at the beginning of the listing.

For the treatment of closed inventory, the **Conflnv** field can have three values:-

1. **default** - title, unitnum and closed until accesdate only is included in the EAD output.
2. **exclude** - no inventory items where access is C are included in the output.
3. **include** - treat closed inventory items the same as other inventory items.

Appendix A HDMS Finding Aid Tables

FindingAid

	Field	Description	Metadata
HDMS	Code	Link Finding Aid to Organisation table	
	Directory	Directory in which to create finding aid files	
General elements	Title	Title of finding aid	DC.title
	Creator	Creator of finding aid	DC.creator
	WebAddress	URL of base directory of finding aid - i.e. filenames are added on creation	DC.identifier
	Copyright	Copyright Statement	DC.rights
	Language	Language of the Finding Aid	DC.language
	PubBy	Publisher	DC.publisher
	PubByHome	URL for publisher	
	PubOn	Published on NOTE - ignore if null	
	PubOnHome	URL for published on	
	PubDate	Date of publication - text field so can be in whatever format is required and include day, month, year as required.	
	PubNumber	ISBN/ISSN etc	
	PrepBy	Listed by	DC.contributor
	PrepByMail	Prepared by mail or URL or null	
	UpdateBy	Updated by	
	UpdateByMail	Updated by mail or URL	
	DateCreated	Date field of first creation date	DC.date.created
DateMod	Date of generation	DC.date.lastmodified	
Content Options	OrderSeriesbyId	Order Series Listing by id if true, by alpha if false	
	OrderProvbyId	Order Provenance by id if true, by alpha if false	
	RemoveSPCodes	Remove Series and Provenance codes and leading zeros	
	ConfSeries	Treatment of Closed Series - default, include, exclude	
	ConfInv	Treatment of Closed Inventory - default, include, exclude	
	OrderInv	Order of Inventory listing - unitnum, item_id, title, date	
	FullProv	Include provenance full note in EAD or create full provenance pages in HTML	
	BoxLoc	Include Box location with Inventory	
	SeriesBoxes	Box numbering within Series - ie. 'Series Box number' is displayed instead of 'Box Number'	
Other	Repository	Repository	
	RepHome	URL for Repository, ignore if null	
	RepReference	Repository's reference id for the collection content	
	Enquiry	Contact person at Repository - name or title	
	EnquiryMail	Email address for contacting Repository	

	Field	Description	Metadata
	Abstract	Abstract	DC.description
	Language	Repository code - for use in EAD ID	
	PrimaryProv_id	reference to primary provenance	

HTML

Field	Description
Picture	Image for top left of pages - e.g. logo, photo
PictureWidth	Width of picture in pixels, default 80
PictureHeight	Height of picture in pixels, default 106
Left	Previous gif
Right	Next gif
Find	Search gif
Colour	Colour of left side bar - ignore if null
ColourWidth	Width of colour bar - default 80 pixels, will ignore if null and let the picture width and browser determine column width.
Background	Optional background image
Home	Filename of home page
Search	Filename of search page
ProvSeriesLinks	Include series list with provenance
How	Text for how to use page
ArcTerms	Text for terms page

EAD

Field	Description
DTDLocation	Location of DTD (on web server)
XMLDTDLocation	Location of XML DTD (on web server)
EntityDeclarations	Any entity declarations Note - making SGML and XML compliant?
Audience	EADHEADER Audience attribute
FindAidStatus	EADHEADER FindAidStatus attribute
CountryCode	EADID Country Code - e.g. AU
OrgCode	EADID Organisation Code
Filename	SGML Filename and EADID Filename content
XMLFilename	XML Filename and EADID Filename content
XSLFile	XSL File name
XSLInc	Include XSL in XML file
ArchdescLevel	Level attribute of archdesc element, i.e. collection, fonds, recordgrp, series, subgrp, subseries, otherlevel

Appendix B Dublin Core Metadata Mappings

NAME	SCHEME	HDMS Field	REF
DC.Title		FindingAid![Title]	4.1
DC.Creator		FindingAid![Creator]	4.2
DC.Subject		publications, guide, records, Australian Rubber Research Academy, archives, records, science, Australia, technology, medicine, ASAP	4.3
DC.Description		FindingAid![Abstract]	4.4
DC.Publisher		FindingAid![Publisher]	4.5
DC.Contributor		FindingAid![PrepBy]	4.6
DC.Date.Created	ISO8601	FindingAid![DateCreated]	4.7
DC.Date.LastModified	ISO8601	FindingAid![DateModified]	4.7
DC.Type		Document	4.8
DC.Format	IMT	text/html, text/sgml, text/xml (updated by the generator according to the type of finding aid chosen)	4.9
DC.Identifier	URL	FindingAid![WebAddress] and HTML![Home] or EAD![Filename] or EAD![XMLFilename]	4.10
DC.Source			4.11
DC.Language	ISO639	en-gb FINDINGAID![LANGUAGE}	4.12
DC.Relation			4.13
DC.Coverage			4.14
DC.Rights		FindingAid![Copyright]	4.15