EAD: The UK Experience

Bill Stockting, the British Library

EAD@10: A Symposium Celebrating the 10th Anniversary of Encoded Archival Description:

San Francisco, 31 August 2008

Introduction

UK archivists started using *beta* version of EAD in 1997 with great success!

A mixed picture though and EAD implementation and use has varied in different sectors of the UK archival community

- The National Archives (TNA) (then the Public Record Office (PRO)
- Local Government – the backbone of UK archival provision!
- Higher Education
- Services offering federated access to catalogues within these sectors

Background

Development of ICT and Internet in 1990's:

- Archivists see potential of ICT tools for access to archives
- Government policy argues for increased access to archives as part of and electronic service delivery in public sector
- Funding available especially from National Lottery

NCA's, *Archives On-line* (1998) sets goal:

a researcher anywhere in the world who has access to the Internet should be able to contact a common gateway, submit a single enquiry and receive an integrated response, listing the relevant source material housed in all UK archive repositories

Standards for Archival Description

- Standards for archival description still not universally accepted in the UK in 1980's but MAD and ISAD(G) codify current practice for full hierarchical finding aids rather than MARC-AMC records

- Different to U.S. experience where EAD seen to influence descriptive practice as set out in DACS

INTERNATIONAL COUNCIL ON ARCHIVES
CONSEIL INTERNATIONAL DES ARCHIVES

ICA

STANDARDS

ISAD(G):
General International Standard Archival Description

Second Edition

Adopted by the Committee on Descriptive Standards
Stockholm, Sweden, 19-22 September 1999

Ottawa 2000

EAD Advantages

EAD had many advantages:

- Designed to work with ISAD(G)

- Based on open technical standards: Initially SGML but early compatibility with XML and related technologies

- Not tied to commercial software and freely available

- Commitment to development and maintenance by international EAD Working Group
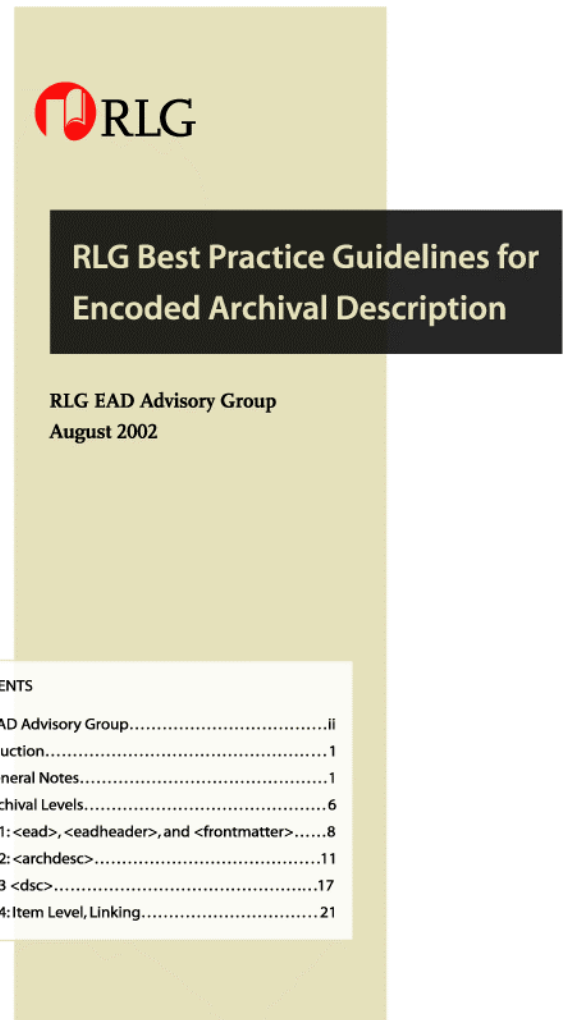
Skills, Knowledge and Experience

Lack of technical, skills, knowledge and experience overcome by:

- Asking the experts

- Working with technical colleagues

- Developing networks and mutual support , nationally (EAD / Data Exchange Group) and internationally

- Training

RLG

RLG Best Practice Guidelines for Encoded Archival Description

RLG EAD Advisory Group
August 2002

CONTENTS

Capturing and Standardising Legacy Metadata

EAD used in process of retro-conversion of legacy finding aids:

- TNA conversion of electronic legacy data using EAD to standardise to ISAD(G) for import to online catalogue

- A2A Programme - conversion of 100,000 finding aids to ISAD(G) using EAD template

- Some university archives use RLG negotiated APEX conversion service

Creating and Managing Standardised Metadata

Different strategies for different needs:

- Some university repositories create EAD files with SGML authoring software and templates

- TNA tests EAD with *Core Executive* pilot but develops hybrid system: SQL server with EAD/XML blobs

- Local authorities use proprietary integrated databases, especially CALM

- Federated services, Archives Hub and Janus, provide tools for creation of EAD by contributors

archives HUB

Online template

Check box if appropriate:
- ☐ **Include subordinate components** within this record (this record will include <dsc> and <c01> elements)
- ☐ This data is **at a subordinate level** (this data forms a <c01> component of a larger record)

Repository Name (e.g. University of Nottingham)

3.1: Identity Statement Area

3.1.1: Reference Code including NCA Repository Code [**required field**]
csa

3.1.2: Title

3.1.3: Dates of Creation   **Normalised Date** YYYY

3.1.5: Extent of Unit of Description

Note: 3.1.4 Level of Description will be generated automatically for this record, with "fonds" as the default.

3.2: Context Area

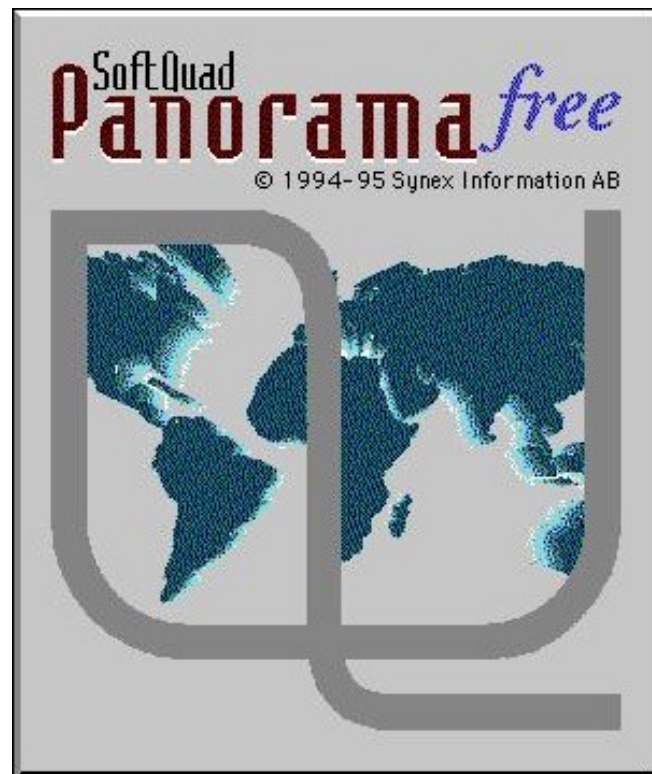**3.2.1: Name of Creator** [**also add manually as Access Point**]

Presenting Metadata

Some present EAD files on the web but:

- HTML browsers can't present SGML so use of plug-ins

- XML and XSL allow transformation of EAD/XML files

- *EAD Cookbook* style-sheets re-used

- Large files an issue:
  - Archives Hub limit to 5mb
  - A2A break up
  - So PROCAT does not use EAD for presentation

Searching Metadata

- Searching of EAD files initially provided by use of *Dynatext* and *Dynaweb* by PRO and some university archives

- Since then different strategies
  - A2A: *TeXtML* now *Autonomy* and XSL
  - Archives Hub: *Cheshire* search engine, Z39.50 and XSL
  - Distributed Archives Hub provides local search interface

Exchanging Metadata

- EAD seen to have a role in data exchange:
  - Government adopts XML
  - NCA interoperability protocol mandates that systems must import and export EAD

- But problem for federated services that no single flavour of EAD the same so provision of tools such Janus' *tidyer*

About Janus | Participating Institutions | Browse and Search | What's New | Feedback | Useful Links | Research Tools

Home > Janus> Webpages> Janus Home Page> About Janus

About Janus contains:
What the "tidyer" does

Search Janus [Go]

Advanced search
Browse catalogues or indexes

What the "tidyer" does

The tidyer changes EAD files in a number of ways. For example:

- it adds a normal form to unitdates which don't have one (this will involve some guesswork);
- it fixes the DOCTYPE declaration;
- it tweaks whitespace e.g. removing spaces around hyphens;
- it fixes several validity problems with EAD produced by Cantab;
- it fixes broken XML encodings e.g. & not & or ¹;
- it canonicalizes unitids into the GBR/1234/ABC format;
- it performs a number of tweaks to punctuation in name authorities to make them conform to local conventions;
- it converts EAD from A2A and Archives Hub, both of which are SGML formats, into something Janus can use.

When the tidyer runs, it provides a short summary of the changes made.

The tidyer saves a copy of the original file with a .bak extension. To find out exactly what the tidyer did you should compare this file with the tidyer's output.

Local conventions re use of punctuation in name authorities

Some local conventions have been adopted governing the use of punctuation in the construction of name authorities. At present, the tidyer will automatically make the following corrections. Contributors are encouraged to note the conventions used and construct name authorities accordingly rather than rely on the tidyer.

1. Removes any full stops before the first opening parenthesis, removes commas around dates, and removes parentheses around words without numbers in them, i.e. **Shakespeare, William. (1564-1616), (playwright)** becomes **Shakespeare, William (1564-1616) playwright**
2. Adds an accent to 'née'.
3. Removes a comma after Sir and before a date, i.e. **Guinness, Sir, Alec, (1914-2000) knight, actor** becomes **Guinness, Sir Alec (1914-2000) knight, actor**
4. Changes a "1945-" style date into "b 1945", i.e. **Clapton, Eric Patrick (1945-) musician** becomes **Clapton, Eric Patrick (b 1945) musician**. And similarly, changes a "-1945" style date into "d 1945", i.e. **Patuone, Eruera Maihi (-1872) Maori chief** becomes **Patuone, Eruera Maihi (d 1872) Maori chief**
5. Changes circa into c, e.g. "**d. circa 1845**" becomes " d c 1845"
6. Turns "**first**" into "**1st**"

W3C XHTML 1.0 | W3C CSS | The webmaster.

Done | Internet | Protected Mode: On | 100%

The Future?

Some areas of debate:

- More complex data model describing record series (not fonds!), creators and functions/activities and their relationships

- User contributed data for 'our' collections and those of Community Archives

- Networking using more distributed model and exchange protocols such as Z39.59 and webs services