

# EAD: The UK Experience

Bill Stocking, 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



ICA

STANDARDS

INTERNATIONAL COUNCIL ON ARCHIVES  
CONSEIL INTERNATIONAL DES ARCHIVES

ISAD(G):  
General International Standard Archival  
Description

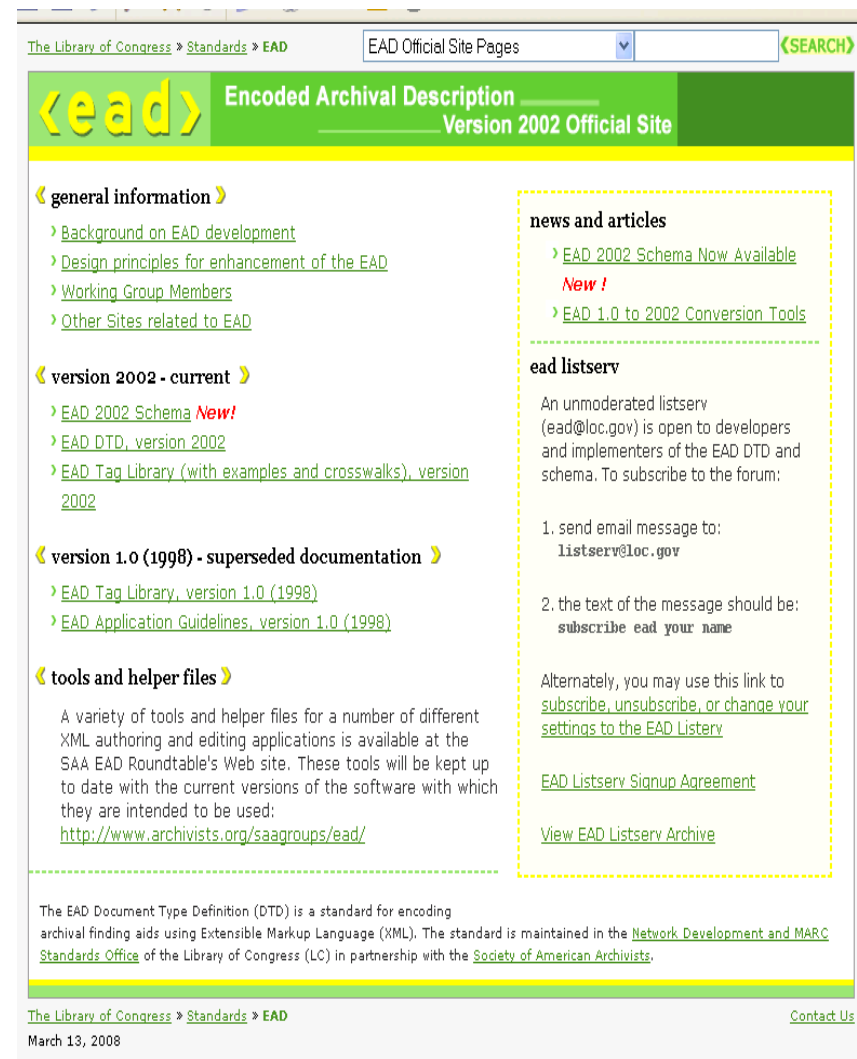
Second Edition

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

# 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



The screenshot shows the 'EAD Official Site Pages' from The Library of Congress. The page features a green header with the 'ead' logo and the text 'Encoded Archival Description Version 2002 Official Site'. A search bar is located in the top right corner.

The main content area is divided into several sections:

- general information**
  - [Background on EAD development](#)
  - [Design principles for enhancement of the EAD](#)
  - [Working Group Members](#)
  - [Other Sites related to EAD](#)
- version 2002 - current**
  - [EAD 2002 Schema \*\*New!\*\*](#)
  - [EAD DTD, version 2002](#)
  - [EAD Tag Library \(with examples and crosswalks\), version 2002](#)
- version 1.0 (1998) - superseded documentation**
  - [EAD Tag Library, version 1.0 \(1998\)](#)
  - [EAD Application Guidelines, version 1.0 \(1998\)](#)
- tools and helper files**

A variety of tools and helper files for a number of different XML authoring and editing applications is available at the SAA EAD Roundtable's Web site. These tools will be kept up to date with the current versions of the software with which they are intended to be used:

<http://www.archivists.org/saagroups/ead/>

On the right side, there is a dashed box containing:

- news and articles**
  - [EAD 2002 Schema Now Available](#)
  - New !**
  - [EAD 1.0 to 2002 Conversion Tools](#)
- ead listserv**

An unmoderated listserv (ead@loc.gov) is open to developers and implementers of the EAD DTD and schema. To subscribe to the forum:

  - send email message to: [listserv@loc.gov](mailto:listserv@loc.gov)
  - the text of the message should be: **subscribe ead your name**

Alternately, you may use this link to [subscribe, unsubscribe, or change your settings to the EAD Listserv](#)

[EAD Listserv Signup Agreement](#)

[View EAD Listserv Archive](#)

At the bottom, a footer section states: 'The EAD Document Type Definition (DTD) is a standard for encoding archival finding aids using Extensible Markup Language (XML). The standard is maintained in the [Network Development and MARC Standards Office](#) of the Library of Congress (LC) in partnership with the [Society of American Archivists](#).'

The footer also includes the text 'The Library of Congress > Standards > EAD', 'March 13, 2008', and a 'Contact Us' link.

# 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 Best Practice Guidelines for Encoded Archival Description

RLG EAD Advisory Group  
August 2002

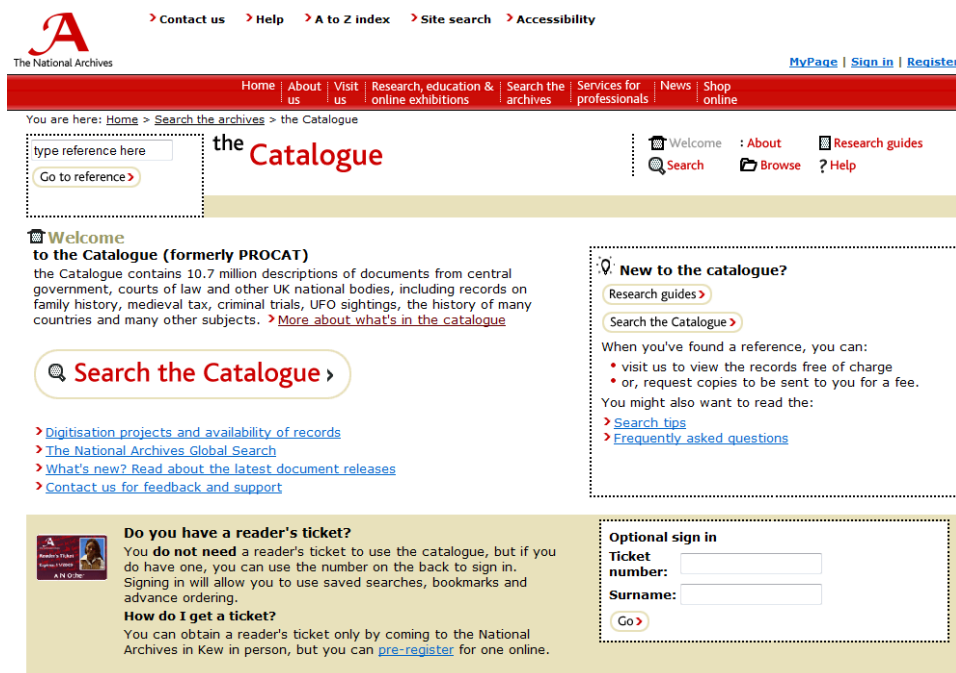
### CONTENTS

RLG EAD Advisory Group.....	ii
Introduction.....	1
I. General Notes.....	1
II. Archival Levels.....	6
Table 1: <ead>, <eadheader>, and <frontmatter>.....	8
Table 2: <archdesc>.....	11
Table 3 <dsc>.....	17
Table 4: Item Level, Linking.....	21

# 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



The screenshot shows the homepage of The National Archives Catalogue. At the top, there is a navigation bar with links: > Contact us, > Help, > A to Z index, > Site search, > Accessibility. Below this is a red banner with links: Home, About us, Visit us, Research, education & online exhibitions, Search the archives, Services for professionals, News, Shop online. A breadcrumb trail reads: You are here: Home > Search the archives > the Catalogue.

The main heading is "the Catalogue". To the left is a search box with the text "type reference here" and a "Go to reference >" button. To the right are links: Welcome, About, Research guides, Search, Browse, Help.

Below the heading is a "Welcome to the Catalogue (formerly PROCAT)" section. It states: "the Catalogue contains 10.7 million descriptions of documents from central government, courts of law and other UK national bodies, including records on family history, medieval tax, criminal trials, UFO sightings, the history of many countries and many other subjects." It includes a link: > More about what's in the catalogue.

Below this is a large "Search the Catalogue >" button. To the right is a "New to the catalogue?" section with links: Research guides >, Search the Catalogue >. It includes text: "When you've found a reference, you can: visit us to view the records free of charge or, request copies to be sent to you for a fee." It also includes links: > Search tips, > Frequently asked questions.

At the bottom left is a "Do you have a reader's ticket?" section. It includes a small image of a ticket and text: "You do not need a reader's ticket to use the catalogue, but if you do have one, you can use the number on the back to sign in. Signing in will allow you to use saved searches, bookmarks and advance ordering. How do I get a ticket? You can obtain a reader's ticket only by coming to the National Archives in Kew in person, but you can pre-register for one online." Below this is a "Optional sign in" section with fields for "Ticket number:" and "Surname:" and a "Go >" button.

# 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



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]

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

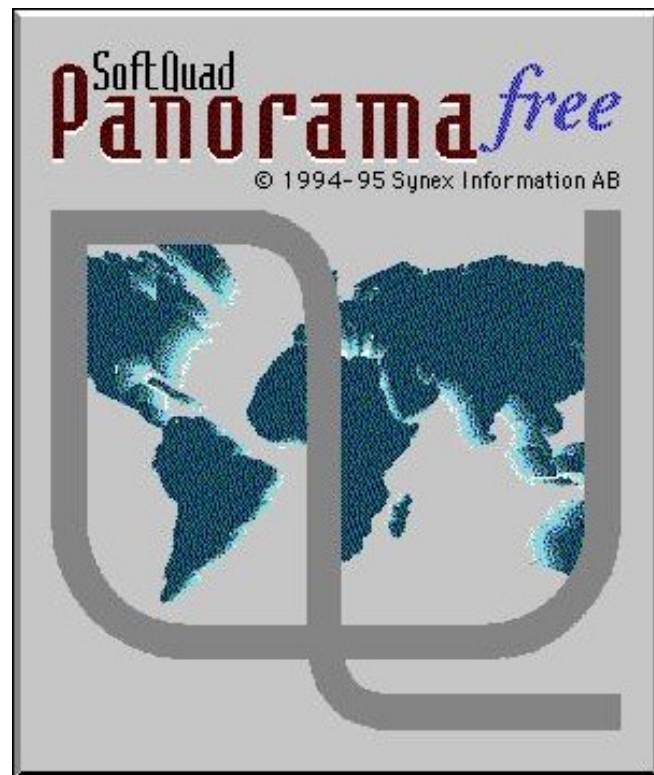
3.2.1: [Name of Creator](#) [also add manually as [Access Point](#)]



# Presenting Metadata

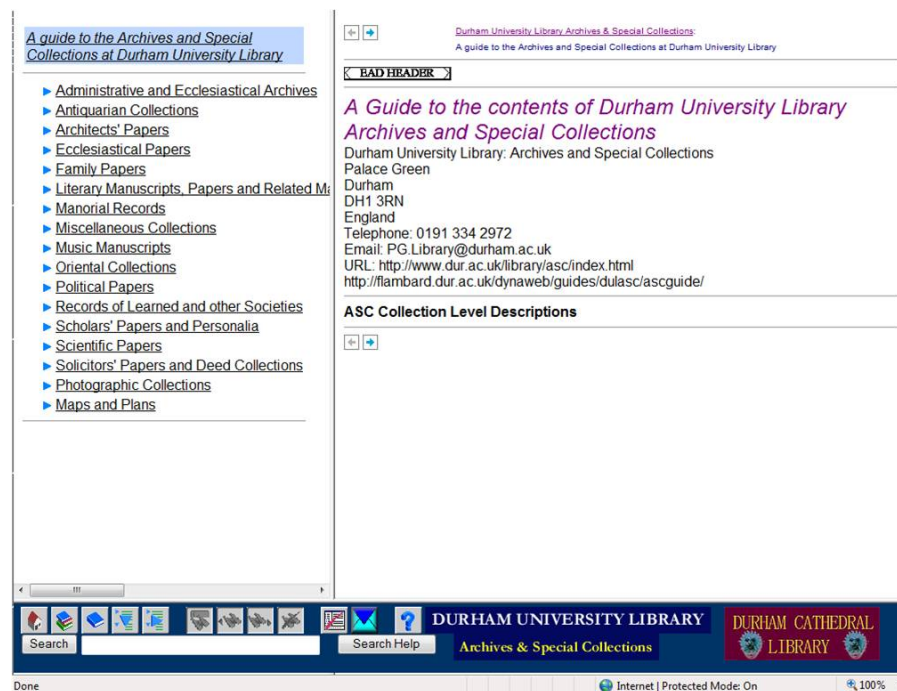
Some present EAD files on the web but:

- 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*

The screenshot shows the Janus website interface. The top navigation bar includes links for 'About Janus', 'Participating Institutions', 'Browse and Search', 'What's New', 'Feedback', 'Useful Links', and 'Research Tools'. The main content area is titled 'What the "tidyer" does' and explains the tool's purpose: 'The tidyer changes EAD files in a number of ways. For example:'. It lists several functions: adding normal forms to unitdates, fixing DOCTYPE declarations, tweaking whitespace, fixing validity problems, fixing broken XML encodings, canonicalizing unitids, and performing tweaks to punctuation in name authorities. A section titled 'Local conventions re use of punctuation in name authorities' provides a list of specific corrections, such as removing full stops before parentheses, adding accents, and standardizing dates and names. The bottom of the page features logos for W3C XHTML 1.0 and CSS, and a 'The webmaster.' link.

# 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

