Weboffice Training – Website Metadata

Introduction

Metadata describes the characteristics of a resource. It can be defined simply as data about data – information describing the content of a resource, and its intellectual property rights.

The application of metadata to web services improves both visibility and accessibility.

The UWA Web Policy requires that the University comply with the *State Records Act 2000 (WA)*. All official web pages must provide metadata in accordance with the University's Metadata Standard¹, and owners of official UWA websites are responsible for ensuring that:

- all official web pages are appraised according to criteria in the University’s Record Keeping Plan²
- published modifications to web pages and metadata are captured in a record management system (TRIM).

What is Metadata?

Metadata is structured data which describes the characteristics of a resource. It shares many similar characteristics to the cataloguing that takes place in libraries, museums and archives. The term “meta” derives from the Greek word denoting a nature of a higher order or more fundamental kind. A metadata record consists of a number of pre-defined elements representing specific attributes of a resource, and each element can have one or more values. Below is an example of a simple metadata record:

<table>
<thead>
<tr>
<th>Element name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Web catalogue</td>
</tr>
<tr>
<td>Creator</td>
<td>Dagnija McAuliffe</td>
</tr>
<tr>
<td>Publisher</td>
<td>University of Queensland Library</td>
</tr>
<tr>
<td>Format</td>
<td>Text/HTML</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Relation</th>
<th>Library Web site</th>
</tr>
</thead>
</table>

Each metadata schema will usually have the following characteristics:

- a limited number of elements
- the name of each element
- the meaning of each element

Typically, the semantics are descriptive of the contents, location, physical attributes, type (e.g. text or image, map or model) and form (e.g. print copy, electronic file). Key metadata elements supporting access to published documents include the originator of a work, its title, when and where it was published and the subject areas it covers. Where the information is issued in analog form, such as print material, additional metadata is provided to assist in the location of the information, e.g. call numbers used in libraries. The resource community may also define some logical grouping of the elements or leave it to the encoding scheme. For example, Dublin Core may provide the core to which extensions may be added.

Some of the most popular metadata schemas include:

- [Dublin Core](#) (DC)
- [AACR2](#) (Anglo-American Cataloguing Rules)
- [GILS](#) (Government Information Locator Service)
- [EAD](#) (Encoded Archives Description)
- [IMS](#) (IMS Global Learning Consortium)
- [AGLS](#) (Australian Government Locator Service)

While the syntax is not strictly part of the metadata schema, the data will be unusable, unless the encoding scheme understands the semantics of the metadata schema. The encoding allows the metadata to be processed by a computer program. Important schemes include:

- [HTML](#) (Hyper-Text Markup Language)
- [SGML](#) (Standard Generalised Markup Language)
- [XML](#) (eXtensible Markup Language)
- [RDF](#) (Resource Description Framework)
- [MARC](#) (MAchine Readable Cataloging)
- [MIME](#) (Multipurpose Internet Mail Extensions)

Metadata may be deployed in a number of ways:

- Embedding the metadata in the Web page by the creator or their agent using META tags in the HTML coding of the page
• As a separate HTML document linked to the resource it describes
• In a database linked to the resource. The records may either have been
directly created within the database or extracted from another source, such as
Web pages.

The simplest method is for Web page creators to add the metadata as part of creating
the page. Creating metadata directly in a database and linking it to the resource, is
growing in popularity as an independent activity to the creation of the resources
themselves. Increasingly, it is being created by an agent or third party, particularly to
develop subject-based gateways.

Reference: Taylor C 2003, An Introduction to Metadata, University of Queensland,
viewed 25 November 2010,

Applying Metadata Schema to UWA Websites as a ‘Backend’ User

Metadata needs to be applied to all standard web pages before they can be made
live. Assets such as text files; RTF files; PDF files; images and forms do not need
metadata, and are generally set up so that metadata schemata cannot be applied to
them.

Before metadata can be applied to a webpage, there must be a metadata schema
active on that page. The standard UWA metadata schemas have been predefined
within Matrix, and are available from the normal asset dropdown menu.

   1. Select the top level node of the website you want to apply the metadata to.
      This will normally be the site itself.
   2. Right click the asset and choose “Metadata Schema” from the dropdown
      menu.

![Choose Metadata Scheme from dropdown menu to see what is currently applied.](image-url)
3. Acquire the locks [Acquire Lock(s)]

4. The “Metadata Schemas” window will launch. If there are schemas already applied, they will be shown (see Figure 1 – above).

5. You will see a dropdown box with the label [New?]; an editable text input box to the right of it, and associated [Change] [Clear] buttons (Figure 2 – page 4).

Press the Release Lock(s) button to release the locks you hold on "University Website Office". [Show Lock Details]

[Release Lock(s)]

Metadata Schemas

<table>
<thead>
<tr>
<th>Applied</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Page design metadata (id: #163040)" /></td>
</tr>
</tbody>
</table>

[New?]

<table>
<thead>
<tr>
<th>Metadata Schema</th>
<th>Change</th>
<th>Clear</th>
</tr>
</thead>
</table>

Cascade Options

[Note] These actions will affect 4250 other assets below "University Website Office"

- [ ] Cascade scheme changes
- [ ] Manually cascade schemas

Figure 2 - Metadata Schemas detail
6. From the New? dropdown menu choose ‘Apply’ (Figure 3 – above).

7. Select the Change button. The hierarchical asset layout on the left will turn purple. Browse to the Metadata node and expand it by left clicking on the associated paddle symbol. You will see a range of Metadata schemas available (Figure 5 – below).

Figure 3 - Choose Apply to add new schema

Figure 4 - Metadata schemas
8. There are two metadata schemas that must be applied to every page. These are the generic “UWA Metadata schema” and the specific “UWA Business Area – XXX” schema.

9. To apply a schema, right click on the associated schema name and select ‘Use Me’ from the pop-up menu. This will copy your selection to the input box.

10. Make sure that the checkbox ‘Cascade schema changes’ is selected (see 5 – above).

11. Repeat steps 7 through 9 for any additional schema that need to be applied. Business Area specific schemas are contained within the Business Area Schemas folder, and you will need to expand this to view, and select, the available schemas (Figure 6 – on page 7).
12. When you are ready to commit these changes, click Commit at bottom right of the screen.

13. This will then apply the chosen Metadata schemas to the asset you selected - normally the site root – plus all its child assets. Depending on the size of the site, this may take a while, and will be accompanied by a HIPO screen.

14. Once it is completed, an Error Report screen will launch (Figure 7 – Page 8), containing messages, warning that the schema cannot be applied to assets such as images, forms etc. This is normal behaviour as we only want to apply the schema to the content nodes. Click ‘Continue >>’ (bottom right of the screen) to continue.
15. This completes the process of applying schemas to web assets. However, you will still be required to complete a number of metadata fields.

UWA metadata schema fields

Introduction

All assets that can be considered a ‘web page’ must have the UWA metadata schema applied to comply with the State Records Act 2000 (WA). Some of the necessary fields are automatically completed, but there are other fields that are Author Defined and need to be completed by the page author before the page can be made live. The UWA metadata schema complies with the Dublin Core Metadata Initiative and more information can be found at:


Edit Author Defined fields

1. Select the required page in the hierarchy tree, and right click. From the resulting pop-up menu, select Metadata. The Metadata screen for the page will launch (8 – page 9).

2. The System Defined and Author Defined fields will be shown.

3. Click on [Acquire Lock(s)] to make the Author Defined fields editable
4. With the page locked for editing, the lower section contains the Author Defined fields.

5. Currently, the fields will be ‘greyed out’ with a tick in the Use Default boxes (Figure 9)

6. Untick the boxes against DC.Subject, DC.Description and DC.Ed-Audience. You can ignore the AGLS.Function box. The Subject and Description fields open for editing, and the Audience field check boxes are made accessible (Figure 10).
**Figure 9 - Page metadata**

<table>
<thead>
<tr>
<th>System Defined</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC.Title</td>
<td>asset_name (default)</td>
</tr>
<tr>
<td>DC.Creator</td>
<td>The University of Western Australia (default)</td>
</tr>
<tr>
<td>DC.Publisher</td>
<td>The University of Western Australia (default)</td>
</tr>
<tr>
<td>DC.Date Created</td>
<td></td>
</tr>
<tr>
<td>DC.Date Modified</td>
<td></td>
</tr>
<tr>
<td>DC.Language</td>
<td>en (default)</td>
</tr>
<tr>
<td>DC.Identifier</td>
<td>asset_url (default)</td>
</tr>
<tr>
<td>GENERATOR</td>
<td>MySource Tools (default)</td>
</tr>
</tbody>
</table>

**Author Defined**

| DC.Subject      | Currently Empty |
| DC.Description  | Currently Empty |
| DC.EducAudience | Currently Empty |

**Reserved**

| DC.EducAudience | Currently Empty |

**Use default value (click to activate)**
Figure 10 - Author defined fields ready for editing

Field – DC.Subject

The field should contain comma-separated keywords that describe or relate to the page in question.

1. Single words only; no hyphens
2. Do not use words similar to ‘and’, ‘or’, ‘from’, ‘in’, etc.
3. Use common abbreviations, but also write them out in full
For example – for the home page of the School of Earth and Environmental Sciences school, earth, environment, geographical, sciences, SEE,

For the Courses – Undergraduate page of the Faculty of Life and Physical Sciences faculty, life, physical, sciences, courses, undergraduate, undergrad, FLPS

For the page for the Graduate Diploma in Mathematical and Statistical Science school, mathematics, statistics, maths, stats, courses, postgraduate, postgrad, graduate, diploma, mathematical, statistical, science, graddip

Field – DC.Description

The field should contain a sentence or phrase that describes the page in question.

- Be inclusive and thorough

For the examples above:

- The School of Earth and Environment at the University of Western Australia
- Courses offered to undergraduate students in the Faculty of Life and Physical Sciences at the University of Western Australia
- Graduate Diploma in Mathematical and Statistical Science offered by the School of Mathematics and Statistics

AGLS.Function

This field can be ignored.

Field – DC-Ed.Audience

Check all boxes that apply to this page. At least one box must be checked.

For example – the home pages of a site would have all boxes ticked. The page for undergraduate courses would have only “Current Students – Undergraduates”, “Future Students – Undergraduates” and “International Students – Undergraduates” ticked.

Business area schema

All assets that can be considered a “web page” must also have the appropriate business area schema applied. This schema allows visitors to easily find the contact email address for a site.

If there is no business area schema, you will need to create a new one.
1. Ask your Information Architect (IA) or Weboffice support to give you the correct contact email to be used for the site

2. In the Hierarchical Asset Map expand the “Metadata” folder and then the “Business area schemas” folder.

3. Use normal methods to clone a copy of any Business Schema inside the folder

4. Right click on the schema you have just created and left click on ‘Details’ on the pop-up menu. Acquire the locks.

5. Change the schema name to “UWA business area – XXX”, where XXX is the name of the site as it appears in the Asset Map e.g “UWA business area – Centre for English Language teaching”

6. In the section “Status”, from the drop down box select “Approve and Make Live”.

7. Click “Commit”.

8. In the “Sections Defined For This Metadata Schema”, click on “Edit Metadata Field”.

![Figure 11 - Choose the business area schema](image-url)
Figure 12 - Set email address in business area metadata

9. In the “Details” section, enter the email address given to you by the weboffice (see Figure 12 – above)

10. Click “Commit”.

11. The schema can now be applied.

Metadata Spreadsheet Procedure (for Weboffice staff)