Access to and Preservation of Texas Electronic State Publications: Metadata for Electronic Publications

The Texas State Library has established an automated system to collect, index and preserve electronic state publications. The Texas Records and Information Locator and Electronic Depository Program (TRAIL/EDP) harvests web sites, retrieving the text, graphics, programming code, style sheets, scripts, sounds and other files and components of each page it encounters within the entity's domain. The TRAIL/EDP search interface address is: www.tsl.texas.gov/trail

Electronic State Publication Requirements

To ensure that publications are appropriately harvested and indexed, the publishing entity must include metadata in its on-line publications, in accordance with Texas Administrative Code, Title 13, §3.4. Metadata describes various aspects of the publication. Special tags in the header of an HTML document indicate the metadata elements. If the encoding of an electronic state publication is not in an HTML or XML format (for example, a Microsoft Word document or an Adobe PDF file), the publisher must create an HTML "wrapper" that contains the metadata elements and links to the resource. The TRAIL/EDP system treats documents that don't have metadata as sub-documents or continuations of the HTML page from which they are linked.

Essential Metadata Elements

<table>
<thead>
<tr>
<th>Name</th>
<th>Alternate Name</th>
<th>Description / Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC.Subject.Keyword</td>
<td>Keywords</td>
<td><em>Significant words</em> from the electronic publication. Pages without this tag are assumed to be part of some other electronic publication.</td>
</tr>
<tr>
<td>DC.Description</td>
<td>Description</td>
<td><em>A narrative description</em> of the electronic publication</td>
</tr>
<tr>
<td>DC.Subject</td>
<td>Subject</td>
<td><em>The topic</em> of the electronic publication, selected from an approved list.</td>
</tr>
<tr>
<td>DC.Type</td>
<td>Type</td>
<td><em>The category</em> of the electronic publication, selected from an approved list. Default value is “Web documents - Undefined.” Publications about state-administered grant programs must be tagged as type “Grants or Funding Opportunities” in order to appear in the TRAIL grant search results.</td>
</tr>
</tbody>
</table>
## Optional Metadata Elements

| Name          | Alternate Name | Description / Content                                                                                                                                 |
|---------------|----------------|----------------------------------------------------------------*************************************************************************************|
| DC.Title      | Title          | Official title of the electronic publication. The TRAIL/EDP system uses an HTML `<Title>` tag if present. The system uses the `<DC.Title>` tag if both tags are present or if the `<Title>` tag is missing. |
| DC.Author     | Author         | Name of the author or creating agency/institution. Default value is the Agency Name from the TRAIL record of Texas government entities.                  |
| DC.Language   | Language       | Language of the electronic publication. Default value is "English."                                                                              |
| DC.Date       | Date           | Date of the electronic publication’s creation. Default value is the file creation date.                                                            |
| DC.Coverage   | Coverage       | Geographic coverage of the electronic publication. Default value is "Texas."                                                                      |
|               |                | Relationships                                                                                                                                       |
|               |                | When necessary, use this tag to define relationships to other print or electronic publications. Content may take the form of "Supercedes print publication 123.AB." |
|               |                | Exceptions                                                                                                                                         |
|               |                | When necessary, use this tag to explain exceptions that prevent archival preservation of a portion of an electronic publication. Content may take the form of "Licensed content from XYZ database has been excluded." |
|               |                | Finding Aids                                                                                                                                       |
|               |                | When necessary, use this tag to identify finding aids or other research tools needed in order to make full use of an electronic publication. Content may take the form of "Index published separately as … ." |

### Standard Metatagging Format

The standard format for adding metadata tags to HTML and XML files is

```html
<META NAME="name" CONTENT="content"/>
```

substituting the appropriate name and content for each metadata element.
Example 1 : HTML web page

A typical web page:

The HTML code showing the metatags in the header area:

```html
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN">

<HTML>
<HEAD>
<TITLE>Texas State Publications Depository Program - Texas State Library</TITLE>
<META NAME="description" CONTENT="This page describes the Texas State Publications Depository Program and provides links (1) to pages with further information about the program's services and products, and (2) to overviews of compliance requirements and forms.">
<META NAME="keywords" CONTENT="Texas State Publications Depository Program">
<META NAME="keywords" CONTENT="Texas State Publications Clearinghouse">
<META NAME="keywords" CONTENT="Texas state documents">
<META NAME="keywords" CONTENT="Texas government publications">
<META NAME="subject" CONTENT="Government information">
<META NAME="subject" CONTENT="State governments">
<META NAME="subject" CONTENT="Libraries">
<META NAME="subject" CONTENT="Executive departments">
<META NAME="subject" CONTENT="Public institutions">
```
Another approach for the same header:

<!DOCTYPE HTML PUBLIC "-/W3C//DTD HTML 3.2//EN">

<html>
<head>
<title>Texas State Publications Depository Program - Texas State Library</title>
<meta name="DC.Description" content="This page describes the Texas State Publications Depository Program and provides links (1) to pages with further information about the program's services and products, and (2) to overviews of compliance requirements and forms.">
<meta name="DC.Subject.Keyword" content="Texas State Publications Depository Program, Texas State Publications Clearinghouse, Texas state documents, Texas government publications">
<meta name="DC.Subject" content="Government information, State governments, Libraries, Executive departments, Public institutions, Government policy, Freedom of information">
<meta name="DC.Type" content="Programs and services">
<meta name="DC.Author" content="State of Texas, Texas State Library and Archives Commission">
<link rel="stylesheet" href="/templates/Styles/master.css">
</head>
[...]

Example 2: Non-HTML Files

This example shows a "wrapper" page (in HTML) that contains links to individual monthly issues of *Texas state publications* in Adobe PDF format files.

The HTML code showing the metatags in the header area:

```html
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN">

<html>
<head>
<title>Texas State Publications - Texas State Library</title>
<meta name="description" content="This page provides links to the monthly issues of Texas State Publications, a monthly index of documents">
<meta name="keywords" content="Texas State Publications monthly index">
<meta name="keywords" content="Texas government documents monthly checklist">
<meta name="subject" content="Government information">
<meta name="subject" content="State governments">
<meta name="subject" content="Libraries">
<meta name="subject" content="Executive departments">
<meta name="subject" content=" Universities and colleges">
<meta name="subject" content="Courts">
<meta name="subject" content="Public institutions">
<meta name="subject" content="Freedom of information">
```

Another approach for the same header:

<!-- DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN" -->

<HTML>
<HEAD>
<TITLE>Texas State Publications - Texas State Library</TITLE>
<META NAME="DC.Description" CONTENT="This page provides links to the monthly issues of Texas State Publications, a monthly index of documents">
<META NAME="DC.Subject.Keyword" CONTENT="Texas State Publications monthly index, Texas government documents monthly checklist">
<META NAME="DC.Subject" CONTENT="Government information, State governments. Libraries, Executive departments, Universities and colleges, Courts, Public institutions, Freedom of information">
<META NAME="DC.Type" CONTENT="Finding aids, indexes, and tracking systems">
<META NAME="DC.Author" CONTENT="State of Texas, Texas State Library and Archives Commission">
<LINK rel="stylesheet" href="/templates/Styles/master.css">
</HEAD>
[...]