

LTER Data Table of Contents (DTC) Data Catalog: A Concise Reference

What is the DTC Data Catalog?

DTC is the “Data table of Contents.” It provides a simple mechanism for creating a searchable LTER-wide data catalog.

What do I need to do to participate in the DTC Data Catalog?

You need to post on your WWW site a file that contains DTC entries for each of the datasets you want to have included. You then need to email jporter@lternet.edu to tell us where your DTC is posted so we can access it.

Each DTC entry must include three mandatory elements: TITLE, LINK and SITE in a special HTML format. In inverse order, SITE is the 3-letter code for your site (e.g. SEV). LINK is the full URL, including the host as well as the directory information (e.g., <http://www.mysite.myuniversity.edu/data/data01.html>), which points to the metadata for that dataset. TITLE is a 1-line (maximum) description of the dataset.

The HTML for a minimum DTC entry would be:

```
<LI><A
href="http://www.myhost.myuniversity.edu/where_metadata_is.html">
The title of the dataset goes here
</A> --VCR
</LI>
```

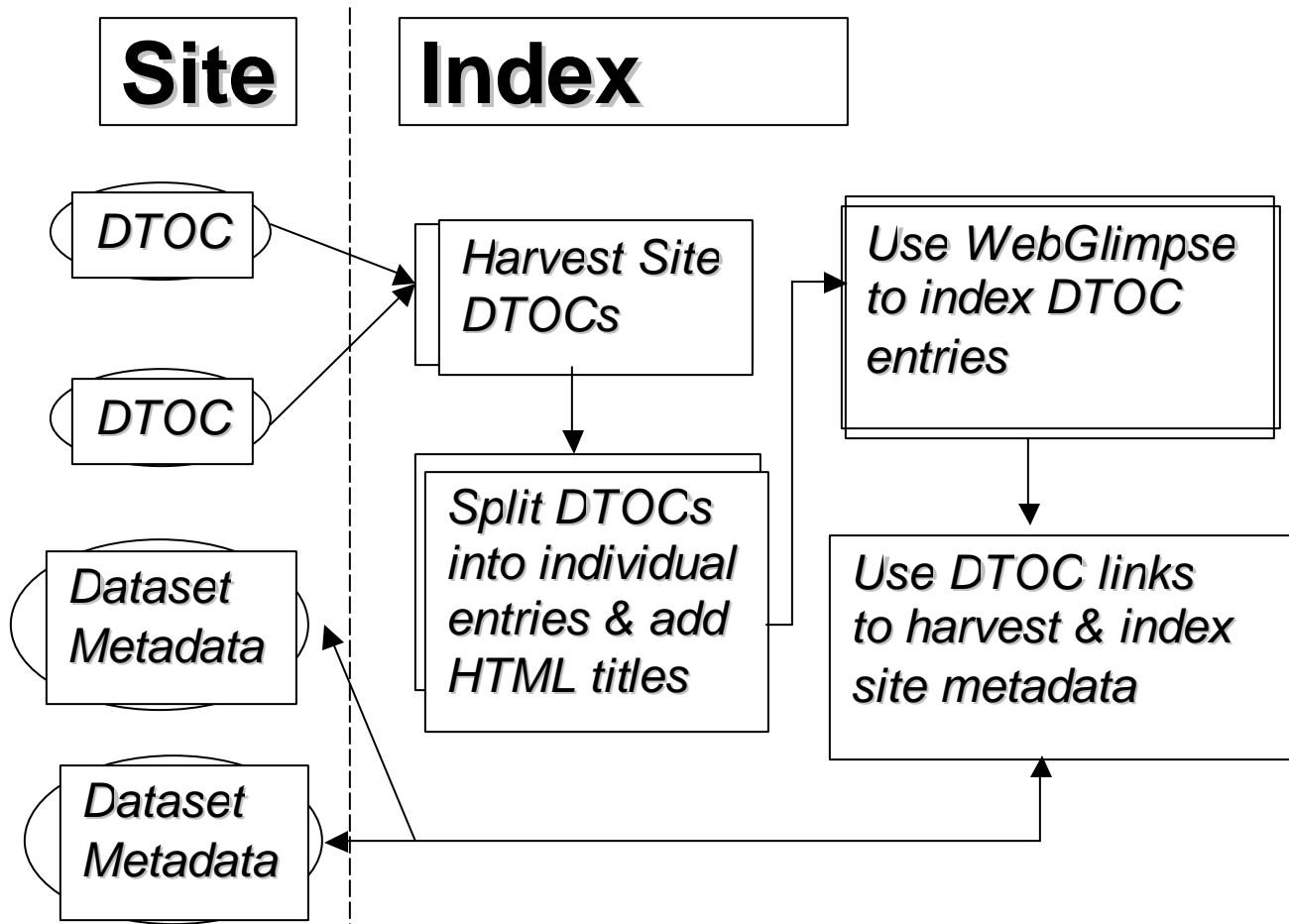
Optionally additional lines can be appended to the DTC to include originators, keywords, dataset identity and other optional information. Each line includes an HTML comment that identifies the type of additional data being provided.

```
<LI><A
href="http://www.myhost.myuniversity.edu/where_metadata_is.html">
The title of the dataset goes here
</A> -- VCR
<BR>---<!-- originators -->Comma separated list of investigators
<BR>---<!-- keywords -->Comma separated list of keywords
<BR>---<!-- datasetID -->Site-specific dataset ID
<BR>---<!-- other -->Other optional information sites want to add
</LI>
```

DTC entries can be prepared manually using a text editor, or linked to your metadata systems so that it is automatically generated (and updated). The choice is up to you. You can view existing entries etc. at the WWW page: <http://www.lternet.edu/DTC> . Contact jporter@lternet.edu with any questions.

How DTOC works

DTOC exploits webcrawler technology to automatically index DTOC entries. In the figure below, on the “Site” side of the figure are things that reside on individual sites. On the “Index” side is the indexing process. A key feature is that the search engine uses the link embedded in the DTOC entry to go back to the site and index the underlying metadata.



Common Problems

The most common problem is not using a full, global, URL in the LINK portion of the DTOC. Partial URLs like “/data/mydata.html” and “mydata.html” lose all their meaning when they are moved from their original location. For this reason use a full URL that includes a machine name like: <http://www.konza.ksu.edu/textfiles/CBC01.html> is required.

A second problem is that the LINK needs to point back to the metadata associated with a particular dataset, not to your main data page, for the indexing to work. Note that we only index the document pointed to by LINK, not documents it subsequently provides links to.