

Chapter 6: Cross-referencing

What's in this chapter?

- Types of cross-references
- The xref element
- Avoiding in-text cross-references to topics
- Cross-referencing topics and external resources
- Cross-referencing elements in a topic
- Links in related-links sections
- Managing cross-references
- Cross-references in titles

There is no greater skill a web professional needs to develop than the ability to create quality links. Many websites do not need more publishing. Rather, they need more linking of content in appropriate task journeys. Linking is a complex skill because it requires you to see the task through your customer's eyes.

Gerry McGovern

Types of cross-references

In DITA, cross-referencing is more than simple hyperlinking, or references to page numbers or section titles. Topics, paragraphs, sections, steps, figures, tables and many other DITA elements can be cross-referenced, along with resources external to the DITA content such as Web addresses, network files, and e-mail addresses.

Cross-references are implemented primarily through the `xref` element and its `href` attribute. Also key to cross-referencing in DITA is the `id` attribute: only elements with an `id` attribute can be cross-referenced.

Cross-referencing features can sometimes be found in unexpected places. For example, the `lq` element includes an `href` attribute, used for storing the URL of the source of the quotation, if applicable. The `source` element in a topic's prolog and the `link` element in the `related-`

links section also have an `href` attribute to store the URL of the original source of the topic content, if applicable.

Cross-references to different elements are resolved, when the collection is processed (published), in different ways. For example, a cross-reference to a `step` element may be rendered as the step number in the output.

The DITA syntax for referencing elements within the same topic is `#[topicid]/[elementid]`.

Avoid generic cross-references to simple paragraphs; instead, cross-references should be to the topic as a whole.

The `xref` element

The cross-reference (`xref`) element allows you to link text within the body of a DITA topic to different *target* resources. Although you can manually specify the text for the cross-reference or link itself, in most cases you should let DITA *calculate* the text based on the type of resource you are linking to. You specify that you want DITA to calculate the cross-reference text by leaving the element empty. For example, if you cross-reference a figure (`fig`) element, the cross-reference text may be automatically generated as `Figure 15`. If you provide text within the `xref` element, that text will be used as the cross-reference text.

Likewise, if you cross-reference a topic, and leave the `xref` element empty, the target topic title will be used as the cross-reference text. For examples of automatically-generated cross-reference text, see [Sample topic - cross-references](#).

The cross-reference (`xref`) element has the following attributes:

- href** nominates the address of the resource to be referenced, such as a topic, a block within a topic, a PDF file, an external Web resource, or an e-mail address.
- type** describes the type of resource being referenced, such as `concept` for a DITA concept topic, `li` for a list item in a list within a topic, `fn` for a footnote within a topic, `fig` for a figure within a topic, or `table` for a table within a topic.

- format** describes the file format of resource being referenced, such as `dita` for a DITA topic, `pdf` for a PDF file, and `html` for a Web resource.
- scope** indicates whether the referenced resource is within the same document as the topic (local), or whether it is external to the document (external).

The syntax of the `href` attribute is illustrated through the following examples:

- Reference to a topic (or the first topic in a composite (*database*) topic): `"file.dita"`
- Reference to a specific topic in a composite (*database*) topic: `"file.dita#topicid"`
- Reference to an element inside a DITA topic: `"file.dita#topicid/elementid"`
- Reference to an element in a ditamap: `"collection.ditamap#map-branch"`
- Reference to an image: `"example.png"`
- Reference to an external resource: `"http://www.example.org"`

Avoiding in-text cross-references to topics

The ditamap's purpose is not limited to specifying the contents of a collection. Automatic inter-topic links are also generated from the ditamap topic hierarchy, and a relationship table section in the ditamap can further extend this automatic linking.

When the same topic appears in a different ditamap, its links will be determined by the hierarchy and relationship table in that different ditamap. In other words, topic-linking is not distributed across many topics, but centralized in the ditamap. This makes maintenance far more efficient.

Cross-references to other topics within paragraphs of a topic can be problematic if the cross-referenced target is not always distributed with the source topic. This can lead to broken links or false assumptions about links being in place.

You should therefore define inter-topic cross-references in the ditamap, using the relationship table (`reltable`) feature.

There is some evidence that linking outside the paragraph text in this way may be more effective.

I've found that the most effective links are written like headings, not part of sentences at all. I've found that putting links in sentences reduces readability and clickability.

Gerry McGovern

Usability tests on the effectiveness of links conducted by Jared Spool support the view that externalized links are more effective than inline, embedded links. Spool's study concluded that:

- Links are less usable when embedded in the text.
- Longer links are more effective than shorter ones.
- People scan for target words - the scent of information.

Cross-referencing topics and external resources

Format of href attribute in cross-references

The resources being referenced in `xref`, `link`, `lq`, `glossref`, and `source` elements are specified in the `href` attribute.

A *URI*, or *Uniform Resource Indicator*, is a standard for referencing resources on the Internet. The similar term *URL*, or *Uniform Resource Locator*, means a specific type of *URI* where the location and the means of retrieving the resource are both incorporated in the address. For example, the *URL* `http://store.scriptorium.com/items/Books/list.htm` nominates the address of the resource and the protocol to use to retrieve it (in this case, *http*). In general, however, most people use the terms interchangeably!

Files on a file server or network drive on a local network can be referenced using the URI convention. Most DITA authoring tools will let you browse for the resource you want to reference, and then format the address in the correct UNC format. For example, the address of a file on a network server must be prefixed with the protocol `smb://`.

You must use (forward) slash characters for path separators. Do not use backslashes. Such file references are URIs in DITA, not file paths. The backslash is an illegal character in URIs. When referencing resources on a network server, prefix the UNC path with `smb://`.

You cannot use an ampersand (&) character in an `href` attribute. If the URL being referenced contains an ampersand character, the `&` character escape should be used to indicate that character.

Links to non-DITA resources

The four most important attributes in the cross-reference (`xref`) element are `href`, `type`, `scope` and `format`.

When linking to non-DITA resources such as PDF files or Web addresses, the `scope` attribute describes whether the linked resource is treated as part of the DITA collection or external to it. This setting will ultimately determine how the HTML output code will be formed during the publishing process, and then how the browser will handle the link. The `format` attribute describes the type of resource being linked to.

The main `scope` attribute values are:

- local** the resource forms part of the content collection (and will typically be opened in the current browser when it is output to a hypertext form)
- external** the resource is managed separately to the content collection (and will typically be opened in a new browser or tab when it is output to a hypertext form)

Note: The other valid `scope` attribute of `peer` is similar to `local`, but is used when the linked resource is not available at build time. It is only relevant when linking to DITA resources.

The valid `format` attribute values for non-DITA resources are:

- html** the non-DITA file format of the linked resource is HTML or XHTML
- pdf** the non-DITA file format of the linked resource is PDF
- txt** the non-DITA file format of the linked resource is plain text
- file type*** the three letter file extension of the non-DITA file format of the linked resource (for example, `ppt` and `odt`)

An example of the code for a local and an external cross-reference link is:

```
<xref href="somefile.html" format="html"
scope="local" />
<xref href="somefile.doc" format="doc"
scope="external" />
<xref href="http://abc.com/intro.htm" format="html"
scope="external" />
```

Linking in a new window

If a topic being referenced in an *xref* is external to the document content, the *xref* element's *scope* attribute should be set to *external* to semantically identify this relationship.

If you want a cross-referenced Web or external resource to open in a new window, code the *xref* as in the example:

```
<xref href=".../somefile.html" format="html"
scope="external" />
```

Note: Opening external target resources in a new window is a processing function, and is therefore dependent upon the publishing tool used and its configuration. This behavior is not part of the DITA standard itself.

Links to glossary terms

In DITA content, the *term* element is used to semantically identify technical or specialist terms, acronyms, abbreviations, and other items of jargon that need to be defined to be understood by some readers. The terms used in a document are normally defined in a separate glossary section.

The glossary entry (*glossentry*) information type can be used to define individual terms; those topics can then be collected into a glossary section in the *ditamap*.

Terms used within normal topic text can be linked, as part of the publishing process, to any corresponding definition of that term in a *glossentry* topic. Linking is accomplished using the *keyref* attribute within the *term* element, and then defining the topic associated with the *keyref* in the *ditamap*, through *glosskey* elements. (The *keyref* attribute was introduced in DITA 1.2.)