

Adobe InDesign 2.0

# Using XML in Adobe InDesign

**E**xtensible Markup Language (XML) is a text file format that lets you reuse content—text, table data, and graphics—in a variety of applications and media. One advantage of using XML is that you can separate form and content. After you add XML tags to your InDesign document, you can export the contents to an XML file. This XML file can then be reformatted for any media, including wireless devices, the Web, eBooks, InDesign® documents, and the printed page.

XML is intended for data exchange, reuse, and publishing on the Web. Like HTML, XML uses elements and structure. However, unlike HTML, XML lets users define a custom set of element tags. Another difference is that XML does not preserve formatting—only content.

## About XML in InDesign

Like an HTML file, an XML text file includes bracketed tags that surround “blocks” of information. For example, to indicate that a sequence of text is a heading in an XML file, a <Heading> tag appears before the text and a </Heading> tag appears after the text. In InDesign, you create tags in the Tags palette and apply them to graphics frames, text frames, and blocks of text within a text frame. Bracketed tags appear when you export the document to XML.

You can export the contents of an InDesign document to an XML file. In addition, you can import the contents of an XML files into an InDesign document.



*InDesign document formatted with XML tags (left) and exported XML file (right)*

Note the following definitions:

**Structure** The sequence and hierarchical relationships of the tagged page items in a document.

**Tag** Labels for page items in your InDesign document. A tag describes the role of the item (such as “Article” or “Sidebar”) and its relationship to other items. In InDesign, tags appear in the Tags palette. XML tags are extensible, not predefined. You can create your own set of tags.

**Note:** Do not confuse XML tags with InDesign tagged text. For more information on InDesign tagged text, which is a different method of exporting and importing InDesign content, see “Tagged-text import options” in the Adobe® InDesign 2.0 User Guide.

**Element** In InDesign, an element is an XML tag that appears in the Structure view window when you tag a page item. Child elements are subordinate to parent elements in the structure. Sibling elements appear on the same hierarchical level.

**Attribute** A string that describes an element using tags or syntax understood in the XML language. Attributes contain a name and a value that can provide information about the element's content, such as when it was created and by whom.

**Page items** Frames, graphics, tables, and text in an InDesign document.

### Installing the plug-in for using XML

Before you can tag your documents or import and export XML, install the XMedia UI plug-in file.

For installation instructions, see the XMediaUI\_Plug-in.pdf file on the Adobe InDesign Cross-Media Publishing Resource CD.

To learn more about working with XML in InDesign, go to <http://www.adobe.com/products/indesign/crossmedia.html>.

### Preparing a document for export

Exporting a document to XML requires four main steps.

**Step 1** Create or load a set of tags. See “Creating and importing XML tags” on page 2.

**Step 2** Apply tags to page items. See “Tagging page items for XML” on page 3.

**Step 3** Use Structure view to revise the levels of hierarchy as needed. See “Structuring documents for XML” on page 5.

**Step 4** Use the Export command to export the structure and content of the InDesign document. See “Exporting a document to XML” on page 7.

### Creating and importing XML tags

Before you tag your document's content for XML export, you need a list of tags in the Tags palette to work with. You can create XML tags from scratch, or you can import them from another source. In most XML workflows, a team member determines the set of tags you'll use—make sure you and your team members use tags consistently.



The Tags palette displays a list of all available tags.


Use any of the following methods to add tags to the Tags palette:

- Create tags using the New Tag command.
- Load tags from an XML file or from another InDesign 2.0 document.
- Import tags (and content) using the Import XML command. See “Importing XML files” on page 8.
- Add tags based on PDF structure tags. See “About using Acrobat structure tags for PDF export” on page 11.

#### To create XML tags:

- 1 Choose Window > Tags to open the Tags palette.

**Note:** If the Tags command does not appear on the Window menu, make sure you copied the Cross-media plug-in properly; see “Installing the plug-in for using XML” on page 2.

- 2 Choose New Tag from the Tags palette menu, or click the New Tag button  at the bottom of the Tags palette.
- 3 Type a name for the tag. InDesign checks to make sure that you include valid XML tag names according to XML standards. If you include a space or an illegal character in the tag name, an alert message appears when you click OK.
- 4 Specify a color for the tag. You can use the same color for multiple tags. The color you select appears when you apply the tag to a frame and choose View > Show Tagged Frames, or when you apply the tag to text within a frame and choose View > Show Tag Markers. Tag colors do not appear in the exported XML file.
- 5 Click OK.

#### To edit a tag:

- 1 Do one of the following:
  - To change only the tag name, double-click a tag name in the Tags palette, and then edit the tag name in its place.
  - To change the tag name or color, double-click the tag color icon. Or, select a tag, and then choose Tag Options from the Tags palette menu.
- 2 Change the name or color of the tag, and then click OK.

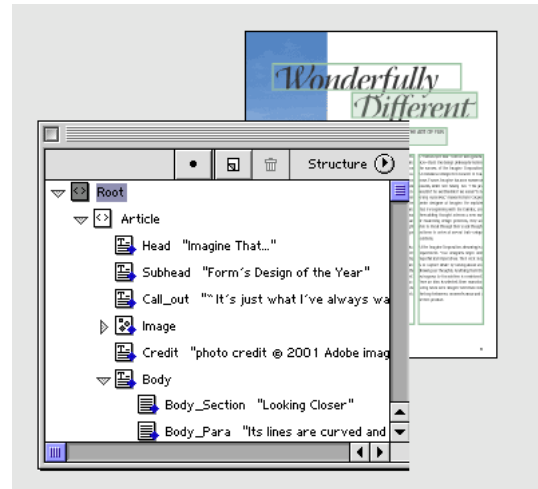
#### To load tags from an XML file or InDesign document:

- 1 Choose Load Tags from the Tags palette menu.
- 2 Select the XML file or InDesign document containing the tags you want to import, and then click Open. Any tags in the selected XML file appear in the Tags palette.

## Tagging page items for XML

You can tag page items such as text frames and graphics frames, including nested frames. You can also tag blocks of text within a text frame to add a level of detail to the XML file. For example, you may want to apply an <Article> tag to a text frame, and then apply specific text tags, such as <Title> and <Body>, to paragraphs within the text frame. (When you tag a frame in a threaded story, all other frames in the story are tagged.)

After you tag items, the elements appear in the Structure view window, which represents the structured content in the document. If you tag a section of text within a frame, the new element appears in Structure view as a child element of the parent frame.





Text markers surround text to which tag has been applied. Elements appear in the Structure view window.

For information on tagging specific elements such as tables and master pages, see “Notes on tagging page items” on page 4.

#### To tag frames:

Do one of the following:

- Drag-and-drop a tag from the Tags palette to a text frame or a graphics frame.
- Using a selection tool, select a text frame or a graphics frame, and then click a tag in the Tags palette.

 If a frame is part of a group or nested within another frame, use the direct-selection tool  to select the frame.

- Using a selection tool, select a text frame or a graphics frame, drag the frame to the Structure view window, and select a tag name from the menu that appears.

**To tag text within a text frame:**

- 1 Make sure that the text frame in which the text appears is tagged. (If the frame isn't tagged when you try to tag selected text within the frame, you are prompted to select a tag for the frame.)
- 2 Do one of the following:
  - Using the type tool, select text within a text frame, and then click a tag in the Tags palette. (If you select all the text within the frame, the tag is applied to the frame, not to the block of text.)
  - Automatically tag paragraphs to which paragraph styles have been applied using the Map Styles to Tags command; see “Mapping styles and tags for XML” on page 10.

**To retag page items:**

Do one of the following:

- Using a selection tool, select a text frame or a graphics frame, and then click a different tag in the Tags palette.
- Using the type tool, place the insertion in a tagged text frame. With Retag selected, click a different tag in the Tags palette to retag the text frame.
- Using the type tool, select the entire block of text to which a tag has been applied, select Retag at the top of the Tags palette, and click a different Tag. (If you select Add Tag at the top of the Tags palette and then click a different tag, a new child element will appear in Structure view.)

**To rename a tag:**

Double-click the tag in the Tags palette, type the new name for the tag, and then press Enter or Return. The new name is applied to all tagged page items.

**Showing or hiding tagged frames and tag markers**

The Show Tagged Frames command displays tagged frames in the color shade determined by the tag color. The Show Tag Markers command displays brackets around tagged blocks of text. The color of the brackets is also determined by the tag color. For information on selecting the tag color, see “Creating and importing XML tags” on page 2.

**To show or hide tagged frames:**

Choose View > Show Tagged Frames or Hide Tagged Frames.



In this example, all page items are tagged except for the subhead below “Wonderfully Different.”


**Note:** If you choose Type > Show Hidden Characters, generic characters for any XML, index, or cross-reference marker will appear.

**To show or hide tag markers:**

Choose View > Show Tag Markers or Hide Tag Markers.

**Notes on tagging page items**

Note the following:

- You cannot tag groups of objects. To tag a page item that's part of a group, use the direct-selection tool  to select the page item.
- You can apply only one tag to a frame.
- Threaded text frames share a single tag, which applies to all text in the thread, including overset text.
- Tables are tagged automatically if the table is in a tagged text frame. You cannot tag a selection within a table. Table tags such as <TR> and <TD> do not appear in Structure view until you export the document to XML.
- When you tag a graphics frame, a reference to the graphic's location is placed in the exported XML file. For more information, see “About importing and exporting graphics for XML” on page 9.

- When you tag selected text, a child element appears in the Structure view window under the text frame tag. When you tag a text frame or graphics frame, an element is added to the appropriate location in Structure view.
- You can tag a master page item, but only one instance of the corresponding element appears in Structure view, regardless of how many times the item appears on document pages. You can manually override and tag master page items that appear on document pages. Overridden master page items appear as separate elements in Structure view.

### Deleting tags and elements

You can delete tags in the Tags palette, and you can delete elements in Structure view. When you delete a tag that has been applied to content, you can specify a different tag that will be applied to the same content instead of the deleted tag. You can also select and delete all unused tags.

When you delete an element in Structure view, both the element and the layout content are deleted, but the frame of the deleted element remains in the layout. If the deleted element is a parent element, the child elements are deleted too.

If you want to delete an element without deleting the content, click Untag in the Tags palette to untag the page item.

#### To delete tags:

- 1 Select the tag in the Tags palette, and then choose Delete Tag from the Tags palette menu.
- 2 If the tag has been applied to content, an alert message appears. Select another tag to replace the deleted tag, and then click OK.

#### To select unused tags:

Choose Select All Unused Tags from the Tags palette menu. If you don't intend to use these tags, you may want to delete them to maintain a clean tag set.

#### To delete an element:

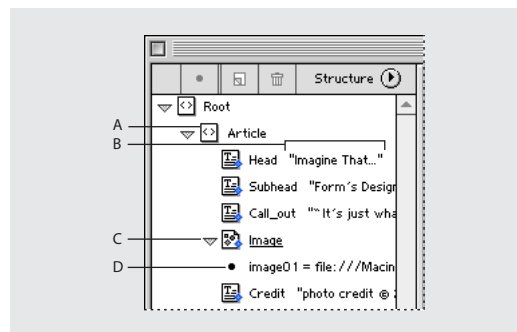
Select the element in the Structure view window, and then choose Delete Element from the Structure pop-up menu in Structure view.

#### To untag page items:

- 1 Select the element in the Structure view window, or select the page item in the document layout.
- 2 Click Untag in the Tags palette.

## Structuring documents for XML


The Structure view window displays a hierarchical tree of tagged page items and imported XML content. Items that appear in the Structure tree are called *elements*. Structure view shows at a glance what elements have been placed on the page and what have not (a blue diamond on an element icon indicates that the element is attached to a page item). You select and move elements in Structure view to define the sequence and hierarchy of page items in the exported XML file.



A. Element name B. Snippet C. Click triangles to expand or collapse elements. D. Attribute



#### To display the Structure view window:

Choose View > Show Structure. If Root is the only element to appear, click the triangle to expand the tags. If no other tags appear, make sure you apply them to page items so that they appear in Structure view.

Click the splitter button  to open the Structure view window to the last saved width, or to close the Structure view window if it's open.








**To resize the Structure view window:**

Do one of the following:

- Position the mouse pointer over the right edge of the Structure view window so that the pointer becomes a double-sided arrow , and then drag left or right to resize the Structure view window.
- Drag the splitter button  left or right to resize the Structure view window.

**Using the Structure view window**

When a tagged page item is selected in the document layout, the corresponding item appears in bold in Structure view. The following icons appear in Structure view:

Icon	Name	Use
	Root element	Each document includes one root document, which can be renamed, but it cannot be moved, deleted, or associated with page item.
	Story element	Represents tagged stories (one or more linked frames).
	Text element	Represents XML-tagged text within a frame
	Graphic element	Represents a tagged frame that includes a placed image. Each graphic element includes an "href" attribute that defines the path or URL to the linked file.
	Unplaced text element	Unplaced text element not yet associated with a page item in the document layout.
	Unplaced graphic element	Unplaced graphic element not yet associated with a page item in the document layout.
	Empty element	An empty frame is associated with this element.
•	Attribute	Represents an attribute.

Any element in Structure view can have *attributes*—optional pieces of information about that element, such as the revision date of an article. Attributes appear only in Structure view.

Story elements include a *text snippet* (the first 32 characters of the story) next to the element. You can hide snippets in Structure view. Table elements do not display snippets.

**To collapse and expand elements:**

Click the triangle next to the element you want to expand or collapse. Hold down Ctrl (Windows) or Command (Mac OS) while clicking the triangle to expand or collapse all elements contained within the current element.

**To show or hide text snippets:**

Choose Structure > Show Text Snippets or Hide Text Snippets. Showing or hiding text snippets sets the default for any new document you create.

**Rearranging elements in Structure view**

You can rearrange the XML structure by dragging elements to different positions in Structure view. Changing the hierarchy affects the structure of exported XML files, but it doesn't affect the position of frames in your layout. However, if you rearrange the order of child elements within a text frame in Structure view, the corresponding text in the document layout is also rearranged. You cannot move the Root element.

**To rearrange the order of elements in Structure view:**

Click the element to select it, and then drag it to a new location within the structure tree. If you want to make an element a child element, select it and drag it to so that the parent element is highlighted, and then release the mouse button.

**Copying and pasting elements**

You can cut or copy elements and then paste them in a new location. When you cut an element, the element and contents are cut to the pasteboard, but the frame remains intact. Be aware that if you paste an element into an untagged frame, the frame will be assigned to the Story tag. You may want to retag the frames to which the Story tag is automatically applied.


**To copy and paste elements:**

- 1 In the Structure view window, select the element or elements you want to copy, and then choose Edit > Copy.
- 2 Select the element where you want the element to appear, and then choose Edit > Paste.

## Adding untagged elements to the Structure view window

To organize your XML content, you may want to create “umbrella” elements in Structure view that are not applied to any page items. For example, you can create a Picture tag that acts as a parent to tags such as Image and Caption.

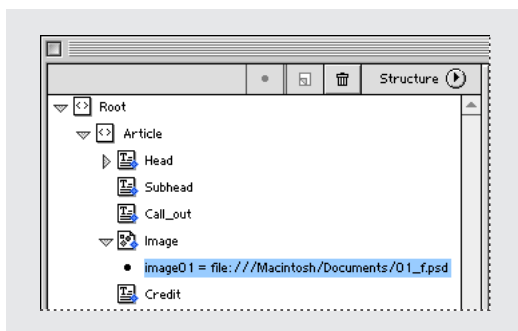
### To add an element to the structure:

- 1 Select the element below which you want the element to appear.
- 2 Select New Element from the Structure pop-up menu in Structure view, or click .
- 3 Select the tag to be used for the element, and then click OK.
- 4 Move the new element within the structure as necessary. Drag tagged elements to the new element to add them as children.

## Creating attributes


An *attribute* describes an element using tags or syntax understood in the XML language. An attribute consists of a name (such as *href* or *keywords*) and a value. When you tag a placed graphic, an *href* attribute appears automatically; it specifies the location of the graphic file. You can edit this href attribute; see “About importing and exporting graphics for XML” on page 9.

You can also add optional attributes to any element. For example, to include keywords for searching in another downstream application, you can create an attribute called *keywords*. You can also create attributes to indicate substituted text, article titles, and other XML content-related information.




An attribute includes a name and a value

### To create an attribute:

- 1 Select an element, and then choose New Attribute from the Structure pop-up menu in Structure view, or click .

- 2 Specify a name and a value, and then click OK.

 *Right-click (Windows) or Control-click (Mac OS) an element in Structure view to display a context menu.*

### To edit an attribute:

- 1 Double-click an attribute, or select an attribute and choose Edit Attribute from the Structure menu in Structure view.

- 2 Edit the attribute as needed, and then click OK.

### To delete an attribute:

Select an attribute, and then choose Delete Attribute from the Structure pop-up menu in Structure view.

### To show or hide attributes:

Choose Show Attributes or Hide Attributes from the Structure pop-up menu in Structure view.

## Exporting a document to XML

After you tag the document’s page items, you’re ready to export a document to XML. You can export the entire set of tagged page items, or you can export only a specified portion. Once the document is exported, you can make additional changes to the XML file using a program such as Adobe GoLive®, or you can import the XML file into an InDesign 2.0 document or another downstream application.

### To export a document to XML:

- 1 Do the following:

- To prepare for export, make sure that tags have been applied to the necessary page items; see “Tagging page items for XML” on page 3.
- If you want to export only a portion of the document, select the element in Structure view where you want exporting to begin.

- 2 Choose File > Export.

- 3 Choose XML for Save as Type (Windows) or Formats (Mac OS).
- 4 Type the name and specify the location of the XML file, and then click Save. The Export XML Options dialog box appears.
- 5 For Comments, type text that will appear in a comment field near the top of the exported XML file.
- 6 To view the exported XML file, select the View XML Using option, and specify the browser or XML editing application in which the exported XML file will appear.
- 7 If you selected an element in Structure view, you can select Export from Selected Element to export only the selected portion of the document. If no element is selected, this option is dimmed.
- 8 For Encoding, select UTF8, UTF16, or Shift-JIS (for Japanese characters). Note that some applications such as Adobe GoLive do not support UTF16.
- 9 Click OK to export the XML file.

**To export tags only:**

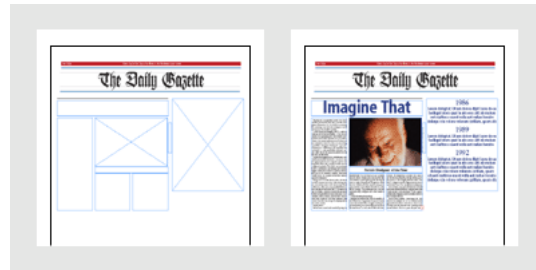
- 1 Choose Save Tags from the Tags palette menu.
- 2 Type a name and specify a location, and then click Save. The list of tags is saved as an XML file.

## Importing XML files

When you import an XML file into an InDesign document, you specify whether imported elements replace existing content or are appended to the contents. Imported XML files contain no text formatting information, because XML separates form and content.

You can flow XML text into any document. If you flow text into a blank document, untagged elements appear in the Structure view window. You can then drag elements from Structure view to the document layout. When you drag and drop elements, the content uses the document's default text and graphics characteristics. You can then reformat this content by mapping tags to styles; see "Mapping tags to styles" on page 10.

You can also flow contents into an InDesign document that acts as a template. The template document contains tagged page items. If you set up your template document properly, the contents of the XML file will flow into the tagged page items of the template document. See "Setting up a document for importing XML tags" on page 9.



*Before and after importing an XML file into an InDesign document*

**To import an XML file and its contents into a blank document:**

- 1 In the InDesign document, choose File > Import XML.
- 2 Select the XML file you want to import, and then click Open.
- 3 To add the imported XML content to the document, drag the elements from the Structure view window into the document layout.

**To import an XML file and its contents into a template document:**

- 1 Open the document in which the XML content will appear.
- 2 Do any of the following:
  - Create a template document for importing XML tags; see "Setting up a document for importing XML tags" on page 9.
  - If the document already contains XML elements, you may want to select an element in Structure view into which the imported structure will flow.
- 3 Choose File > Import XML.
- 4 Select the XML file you want to import.



5 Do one of the following:

- Select Replace Content to replace existing tagged page items; see “Setting up a document for importing XML tags” on page 9.
- Select Append Content to create new elements at the end of the structure.

6 Select Import into Selected Element to import XML content into the element selected in Structure view. If this option is not selected, imported elements will appear under the root element.

7 Click Open.

### Setting up a document for importing XML tags


When you select Replace Content in the Import XML dialog box, the content flows into the elements with matching tags. Replacement starts where the root of the XML file matches the tag of the selected element.

The key is to be rigorous about setting up your tagging structure, so that tag names and values are used consistently in your XML files and your InDesign templates. When you set up the document into which the XML file is imported, be aware that the order in which elements are inserted into a document is determined by the order of the XML file, not by the InDesign file. For example, if you have an InDesign document with elements in Tag2, Tag1 order and the XML file lists these elements in Tag1, Tag2 order, only Tag1 will appear in the InDesign document; Tag2 will appear in the Structure view window, but will not be flowed into the document.

If you want XML elements to replace existing content, you can create a template that includes placeholder text with appropriate styles.

When creating a template, do the following:

- Create text and graphics placeholders in an InDesign document, and then tag them using the same tag names used in the XML file.

 *To make sure you're using the same tag names, you may want to load the tags from the XML file; see “Creating and importing XML tags” on page 2.*

- Make sure the tagged page items appear in the same order as they appear in the XML file. Compare the order of the tags in the XML file with the order of the elements in the document's Structure view. If the order is different, rearrange the elements in Structure view so that they coincide with the order of the XML file's tag order.
- If you want text to retain its formatting during XML import, apply paragraph styles to the tagged placeholder text within a frame, or to the text frame itself. The imported XML text will retain the style's formatting, but the style will no longer be applied to the imported text.
- If individual paragraphs are tagged in the imported XML file, you may want to tag paragraphs within text frames using the same tags and using the same order.
- Make sure tables include the same number of rows and columns as the tables in the imported XML file. Imported table data will replace the contents of the placeholder table, but the table's structural formatting will remain intact.
- You can import XML content into placeholders on master pages. However, automatic page number characters used in headers and footers are not preserved.

### About importing and exporting graphics for XML

When you tag a placed image, InDesign automatically creates an attribute called “href” that defines the location of the image file. You can edit this attribute to specify a new location of the image. For example, your image files may be on your hard drive, but when you transfer the XML and associated image files, these href attributes will no longer be valid on someone else's computer. You can edit the attribute to add relative paths so that others will have access to these files.

By the default, an href attribute may look like the following:

```
href=file:///C:/Photos/Meadow.psd
```

If the image files will appear in the same folder as the exported XML file, edit the attribute as follows:

```
href=file:Meadow.psd
```

If the image files will appear in a subfolder in the folder in which the XML file is located, edit the attribute as follows:

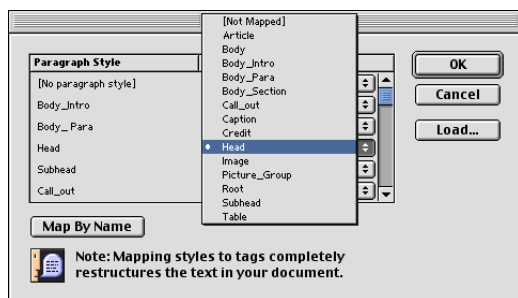
```
href=file://Photos/Meadow.psd
```

#### To edit an href attribute:

- 1 In the Structure view window, double-click the attribute (you may need to click the triangle icon next to the image element to display the attribute).
- 2 For Value, type the new path of the image, and then click OK.

## Mapping styles and tags for XML

You can save time by mapping styles to tags before you export to XML, or by mapping tags to styles after you import an XML file. You can edit the mappings at any time.



Mapping styles to tags

### Mapping styles to tags

While setting up your InDesign document for XML export, you can apply XML tags to paragraphs that have been formatted with paragraphs. For example, if you've used a Heading style and a Body style to format paragraphs in a text frame, you can use the Map Styles to Tags dialog box to map these paragraph styles to equivalent XML tags. When you click OK, paragraphs that are formatted with the mapped styles are also tagged for XML.

If you've already added tags to your document, you may not want to map styles to tags. Mapping styles to tags before export removes any existing tags within tagged text frames; the text becomes tagged as dictated by the current map set.

#### To map styles to tags for export:

- 1 Set up the necessary styles in your document.
- 2 Choose Map Styles to Tags from the Tags palette menu.
- 3 To match style names to tag names, click Map by Name. Styles that are named identically to tag names are selected in the dialog box, not created. Map by Name is case-sensitive, so *H1* and *h1* would be treated as different tags.
- 4 For each style in the current document, specify the tag that you want it to be mapped to.
- 5 To import styles from another document, choose Load, specify the file, and then click OK.
- 6 Click OK.

### Mapping tags to styles

When you map tags to styles, you make sure that tagged paragraphs are also formatted with paragraph styles. You can apply the same style to different tags. For example, you can map Head1 and Head2 paragraph styles to the H1 tag.

#### To map tags to styles for import:

- 1 Set up the necessary tags in your document; see "Creating and importing XML tags" on page 2.
- 2 Choose Map Tags to Styles from Tags palette menu.
- 3 For each tag in the current document, specify the paragraph style that you want it to be mapped to.
- 4 To match style names to tag names, click Map by Name. Styles that are named identically to tag names are selected in the dialog box, not created. Map by Name is case-sensitive, so *H1* and *h1* would be treated as different tags.
- 5 To import styles from another document, choose Load, specify the file, and then click OK.
- 6 Click OK.

## About using Acrobat structure tags for PDF export

Adobe Acrobat includes a set of structure tags, such as Article, Figure, and Artifact, for creating structured PDF files. If you select Include eBooks Tags when you export a document to PDF, these tags are applied to the document automatically. However, you can use Structure view to add additional Acrobat tags or to fine-tune existing tags.

By applying Acrobat tags to your document, you can create a re-flowable PDF file for viewing on handheld devices and other media. For example, if you tag a page item as Artifact, the page item will not be displayed in Reflow view on some handheld devices.

### Adding structure tags to the Tags palette

When you choose the Add Untagged Items command, InDesign adds Acrobat structure tags to the Tags palette, and applies the Story and Figure tags to certain untagged page items. The Story tag is applied to any untagged text frames, and the Figure tag is applied to any untagged graphics. You can then manually apply other tags to sections of text. However, tagging page items does not guarantee that the items will be structured accordingly in the exported PDF file. For more information, see “Applying Acrobat tags to the InDesign document” on page 11.

### To add PDF structure tags to the Tags palette:

- 1 Choose View > Show Structure to display the Structure view window.
- 2 Choose Add Untagged Items from the Structure pop-up menu in Structure view.

## Applying Acrobat tags to the InDesign document

Note the following suggested uses for each imported tag:

**Story** Use this tag for text frames. The Story tag is applied to all untagged text frames when you choose Add Untagged Items.

**Figure** Use this tag for placed graphics. The Figure tag is applied to all untagged graphics placed in your document when you choose Add Untagged Items.

**Artifact** The Artifact tag lets you hide page items when viewing the exported PDF file in Reflow view. This is especially useful for viewing PDF files on handheld device or in other PDF readers.

**Paragraph tags (P, H, H1–H6)** These tags have no effect on the exported PDF text when viewed in Reflow view. However, they can be useful in some situations when exporting a PDF file to HTML format.

