

Using XML

In this lesson, you'll learn how to do the following:

- Import XML tags.
- Apply XML tags.
- Map styles and tags.
- Use the Structure pane.
- Import XML.
- Export XML.

Getting started

In this lesson, you'll take a completed InDesign layout and apply XML tags to the document. You'll then confirm the structure of the document, export the contents as XML, and then import the XML into another InDesign layout. Before you begin, you'll need to restore the default preferences for Adobe InDesign.

1 To ensure that the tools and palettes function exactly as described in this lesson, delete or deactivate (by renaming) the InDesign Defaults file and the InDesign SavedData file. See “Restoring default preferences” on page 2.

2 Start Adobe InDesign CS2.

To begin working, you'll open an existing InDesign document.

Note: If you have not already copied the resource files for this lesson onto your hard disk from the Lesson_13 folder from the Adobe InDesign CS2 Classroom in a Book CD, do so now. See “Copying the Classroom in a Book files” on page 4.

3 Choose File > Open, and open the 13_a.indd file in the Lesson_13 folder, on your hard disk.

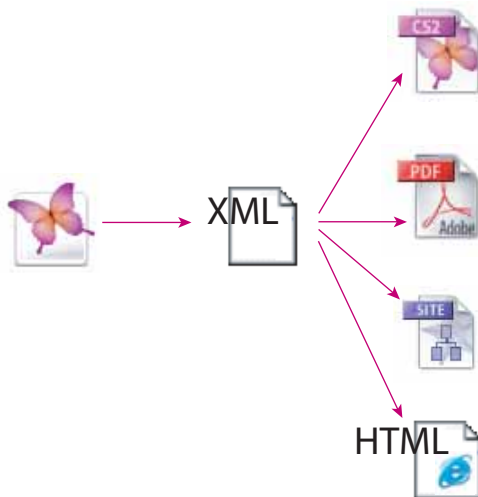
- 4 Choose File > Save As, rename the file **ski_brochure.indd**, and save it in the Lesson_13 folder.

About XML

eXtensible Markup Language (XML) is used to distribute content, including text and images, to multiple destinations. XML separates content from its appearance on a page or in a layout. XML files use tags that identify the content in the file. Because the content of an XML file is identified with tags, it can be repurposed for distribution in print, online, or in some other format, such as PDF.

XML files are not dependent upon a specific layout, and can be formatted to match the needs of the viewer or viewing device. For example, an XML version of a typical brochure might be formatted to display as a traditional vertical layout when printed. The same information can be converted to a horizontal layout for on-screen viewing as a PDF. The same content can also be converted to HTML and placed onto a web page.

While XML allows for data to be presented in a variety of formats, it also has advantages for revising and customizing print layouts. When InDesign documents are converted to XML, you can easily extract some or all of the data to re-use in other InDesign layouts. For example, a single XML file can be the source for a sales brochure, a price list, and a catalog.




If you design web sites using Adobe GoLive or other programs, you may be familiar with Hyper Text Markup Language (HTML). While XML and HTML have similar names, and both markup languages include tags, they are actually quite different. HTML tags describe the appearance of the content in an HTML file, and how it should be formatted. Conversely, XML tags describe the content itself and do not describe formatting. With XML, formatting decisions occur when the XML is placed into a specific layout.

To use the XML features of InDesign CS2, you do not need extensive XML knowledge. To learn more about the XML capabilities of InDesign CS2, choose Help > InDesign Help. Then select XML in the InDesign Help Contents.

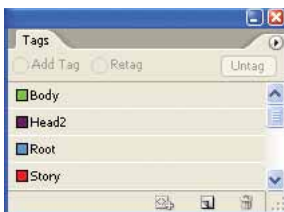
Viewing XML tags

InDesign CS2 uses XML tags to identify the type of content that is, or will be, placed on a page. These tags are used when XML is either imported or exported from an InDesign layout. Frames can be tagged, or specified, to contain certain types of text, such as stories or headlines. Frames can also be tagged to contain graphics. The same tags may be used multiple times within a document. For example, a document may contain multiple headlines, and each headline may contain the same tag name identifying it as a headline.

 *If you regularly work with XML, you may be more familiar with XML elements. Tags applied using InDesign CS2 identify the specific occurrence of an XML element.*

In this exercise you'll examine some existing XML tags that have been applied to an InDesign CS2 layout.

- 1 Choose Window > Tags to open the Tags palette. The Tags palette opens, displaying a list of XML tags that have previously been created with this document.



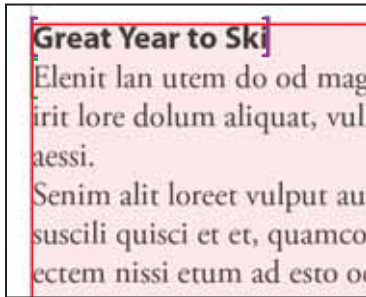
- 2 Choose the Selection tool (⬠) and click the text frame containing the text “Snow and Ski Report” along the top of the document. Notice that the Title tag in the tags palette becomes highlighted as this frame is selected. The highlighting indicates that the Title tag is applied to this text frame.
- 3 Choose Edit > Deselect all.
- 4 Choose the Type tool (T) and click in the headline text “Great Year to Ski.” In the tags palette, the tag Head2 becomes highlighted, indicating that this text also has an XML tag applied to it.
- 5 Choose View > Structure > Show Tagged Frames. The two text frames we’ve examined display a colored border and background.


The border and background color corresponds to the color of the XML tags that are applied to the frames. Notice that the Title tag is Magenta and this is the color that is displayed around the title of the publication.



- 6 Choose the Zoom tool (🔍) and click and drag a box around the text “Great Year to Ski” along with several lines of type beneath this text. The magnification of this area of the page increases.

7 Choose View > Structure > Hide Tag Markers. The XML tag markers at the start and end of the Great Year to Ski headline disappear. Choose View > Structure > Show Tag Markers. Markers are displayed to the left of the word Great and the right of the word Ski. The markers are purple, which corresponds to the color of the Head2 tag in the Tags palette.



 Tag markers are also visible in the Story Editor.

8 Double-click the Hand tool () to view the entire page.

Note: The tag indicators help identify which objects have been tagged, and these markers display only on screen, not when a document is printed.

Importing and applying XML tags

You can create XML tags within your InDesign CS2 layout, or you can import them from another InDesign CS2 document or a Document Type Definition (DTD). In this exercise, you will import several new tags, and then apply the tags to objects in the layout.

1 Select the Type tool (T) and click anywhere in the text of the sidebar titled “Super Skiing” along the right side of the document. No tags in the Tags palette are highlighted, as tags have not been applied to this text.

When the cursor is placed within tagged text, or when a tagged object is selected, the tag applied to the text or object is highlighted in the Tags palette. Because this text is not yet tagged, no tag name is highlighted. You will import several new tags, and then apply them.

- 2 Choose Edit > Deselect All.
- 3 In the Tags palette, choose Load Tags from the palette menu. Navigate to the Lesson_13 folder and choose the tags.xml file, then click Open. InDesign adds two new tags into the Tags palette: Head1 and Image. You will apply these tags to objects and text in your layout.

Note: The names of the tags must exactly match those used in your XML workflow. XML is very precise, causing head1 and Head1 to be considered different tags.

- 4 In the tools palette, choose the Selection tool (⬅), click the text frame containing the headline text “Family Fun,” and click the Head1 tag in the Tags palette. The Head1 tag is now applied to this text frame and all its contents. The frame containing the Family Fun text becomes blue, reflecting the tag that has been applied to it.



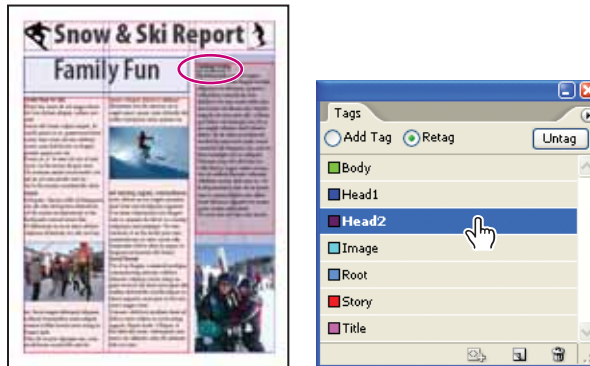
- 5 In the Tools palette, select the Type tool (T). Click and drag to select the headline “Good Snow” in the second column. Click the Head2 tag in the Tags palette. The tag is applied to the selected text. Brackets appear at the start and end of the selected text, identifying that the text is tagged. The color of the brackets corresponds to the color of the tag. If necessary, choose View > Structure > Show Tag Markers to display the text tags.

Note: You may need to deselect the text to clearly see the tag markers that are applied to it. Choose Edit > Deselect All to deselect the text.

- 6 Click and drag to select the body text that follows the “Good Snow” headline. Be sure to select the text in the final two paragraphs of this story, at the bottom of the second column. While the text is selected, click the Body tag in the Tags palette.

7 If necessary, scroll to view the sidebar in the far right column of the page. Continuing to work with the Type tool, click and drag to select the “Super Skiing” headline, and then click the Head2 tag to apply the tag to this text.

As the Head2 tag is applied, tag markers are displayed at the start and end of the tagged text. The frame also changes colors, as a tag was applied to it at the same time you applied a tag to the headline. This is because a frame must have a tag if any of the contents are also tagged. The tag applied to the frame may be different from the tag applied to the text.



8 Continuing to work in the sidebar, click and drag to select the remaining text in the third column. Be sure to select all the text that follows the “Super Skiing” headline, without selecting the headline itself. With the text selected, click the Body tag in the Tags palette.

You have completed tagging all the text in the layout.

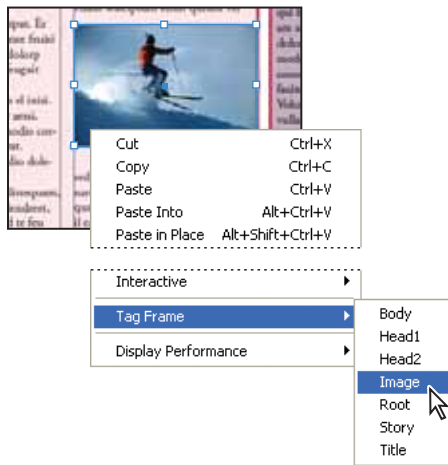
9 Choose Edit > Deselect All.

10 Choose File > Save.

Tagging images

- 1 In the Tools palette, choose the Selection tool (⌘).
- 2 Click the image of the snow volleyball players in the far left column to select it. In the Tags palette, click the Image tag so that it is applied to this graphic.
- 3 Click the image of the downhill skier in the second column to select it. Right-click (Windows) or Ctrl+click (Mac OS) and choose Tag Frame > Image from the contextual menu that is displayed.

You can use the contextual menus to apply tags to frames as well as text.



- 4 Repeating the process from the previous step, click to select the image of the two skiers in the third column, and then right-click (Windows) or Ctrl+click (Mac OS) and choose Tag Frame > Image from the contextual menu.

You have completed the process of tagging all the frames for content that will be exported as XML.

Viewing and organizing structure

Before exporting the document as XML, you want to confirm that the structure of the XML matches the hierarchy of your layout.

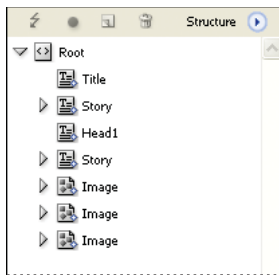
1 Choose View > Structure > Show Structure. The Structure pane is displayed.

The Structure pane displays the tags that have been applied to objects in the layout, in the sequence in which they were applied.

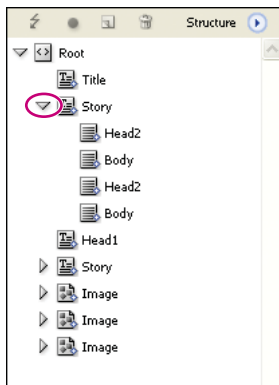
Note: You may need to move your Tools palette to view all the contents of the Structure pane.

2 In the Structure pane, click the triangle (▷) to the left of the Root tag. The Root tag is the highest-level tag; all other tagged objects appear beneath the Root tag. All documents must have at least one Root tag.

The Root tag can be renamed if your workflow uses a different name for its top-level tag.



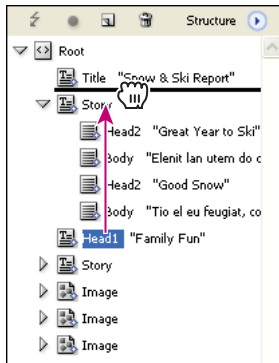
3 Click the triangle to the left of the top Story tag. All the structured elements under this tag are displayed, including both Head2 tags and both Body tags.



You can see the tags that you applied, as well as tags that had been previously applied to the document. The Head2 and Body tags are children of the Story tag, which was previously applied for you. The tags are listed in the order they were applied to objects on the page, and should match the relationship of objects in the layout. For example, all related text is typically placed under the same parent tag. In this case, the Head2 and Body tags are related, so they are placed together under the Story tag. Each Head2 tag is placed before each Body tag, just as the headlines are before the body text in the layout.

4 In the Structure pane, click the menu in the upper right corner and choose Show Text Snippets. The Structure pane now shows the first portion of the text to which each tag is related.

5 In the Structure pane, click the Head1 Tag, which displays the text snippet “Family Fun.” In the layout this is the highest-level headline, placed immediately beneath the title. In the Structure pane, this tag is positioned between two Story tags. As a primary headline, it should be one of the top-most items in the structure. Because InDesign adds items to the Structure pane in the order in which they are tagged, it is incorrectly positioned. Click and drag the Head1 tag up so that it is positioned under the Title tag. Release the Head1 tag when a black line appears under the Title tag. The position of the tag in the Structure pane now reflects the position of this headline in the layout.



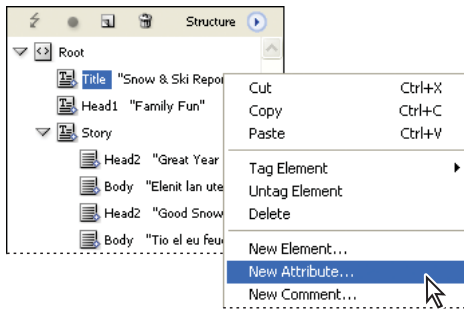
Note: InDesign CS2 also supports using DTDs to validate the structure. For more information about using DTDs to validate structure see InDesign Help.

6 Choose File > Save.

Viewing and applying attributes

Tagged objects may also include additional information about the objects. This information which describes the tagged object is known as an attribute.

- 1 In the Structure pane, click the triangle to the left of the top Image tag. InDesign displays the attribute icon (●) followed by the location of the image file.
- 2 Continuing to work in the Structure pane, click to select the Title tag with the text snippet “Ski & Snow Report.” Right-click (Windows) or Ctrl+click (Mac OS) the Title tag and choose New Attribute. The New Attribute window opens.



- 3 In the New Attribute window, for Name enter **Issue_Number**. Press the Tab key, and for Value enter **12**. Click OK and the attribute is added to the Title tag. Attributes can be used to provide additional descriptive information such as the publication date or copyright status.



- 4 Choose File > Save to save the document.

Exporting XML

After applying XML tags to the text and graphics in the InDesign layout, you can create an XML file.

- 1 Choose File > Export. The Export window opens.
- 2 If necessary, in the Export window, navigate to the Lesson_13 folder, then choose XML from the Save as type drop-down menu (Windows) or Format drop-down menu (Mac OS). Enter the file name **snow.xml**. Click the Save button and the Export XML window opens.



- 3 In the Export XML window, click the Images tab. Confirm that all the options are unchecked.

When exporting XML from InDesign CS2, you can have images from the layout copied or moved into the same folder as the XML file by selecting the options under the Copy to Images Sub-Folder section of this window. This allows for images that have been tagged to be re-used in other print or web layouts. Because you will be working with the XML file using the same computer that extracted the XML, none of these options is necessary.



- 4 Click the Export button. The XML file is generated. Keep this document open, as you will return to work in it later in this lesson.

Importing XML

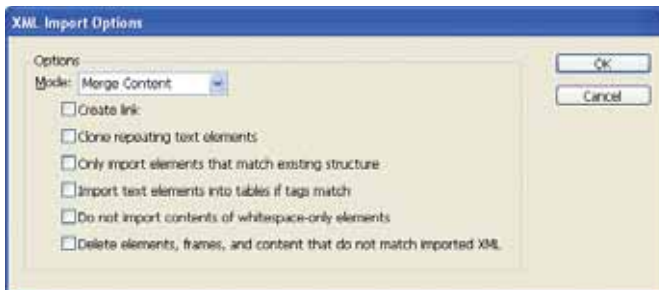
Next you will open an existing InDesign layout that already includes XML tags, and you will import XML that was created in the previous exercise.

- 1 Choose File > Open, and open the 13_b.indd file in the Lesson_13 folder, on your hard disk.
- 2 Choose File > Save As, rename the file **online_brochure.indd**, and save it in the Lesson_13 folder.
- 3 Choose View > Structure > Show structure. The Structure pane is displayed.
- 4 In the Structure pane, click the triangle to the left of the Root tag. All the tags that have been applied to objects in this document are displayed.

Note: The top-level structure of the online brochure matches the structure used in the original print layout.

- 5 Choose File > Import XML. The Import XML window opens.

In the Import XML window, choose Show XML Import Options and click to select the Merge Content radio button. Click to select the snow.xml file, then click the Open button. The XML Import Options window opens.



- 6 In the XML Import Options window, confirm the Mode drop-down menu is set to Merge content and deselect all the check boxes so that all the options are unchecked, and click OK. The XML is imported into the horizontal layout designed for on-screen viewing.

All the text and graphics which were tagged in the original document are flowed into their new layout. The text will be formatted in the next section of this lesson.

Appending versus merging XML

When importing XML, you can either append or merge the XML content into your document. Appending adds the new XML content to your document, leaving the existing structure and content unchanged. Merging replaces existing content and, depending on the options you select, adds new content where it finds no equivalent elements. You merge XML into a document in the following situations:

- The document contains placeholder frames that you want to fill with the incoming XML file.
- The document contains XML that you want to replace with the incoming XML file.
- The document doesn't contain any XML structure, and you want InDesign to replace the default root with the root of the incoming XML file.

XML import options

When importing XML files using the Merge Content option, the XML Import Options dialog box has the following options:

Create link—InDesign creates a link to the XML file and lists the XML file in the Links palette. If the XML file changes, the modified link icon is displayed and you can update the XML file just as you would any other linked item in your layout.

Clone repeating text elements—Replicates the formatting applied to tagged placeholder text for any repeating content. The tags and structure of the XML content must match the tags and structure in the placeholder text. Use this option when importing content with multiple items that all have the same structure, such as a list of names and addresses, or products and prices.

Only import elements that match existing structure—Filters the imported XML content to include only elements that match the tags and structure in the Structure pane.

Import text elements into tables if tags match—Imports elements into a table if the tags match the tags applied to the placeholder table and its cells. Use this option to place database records into a table, for example, when generating price lists or inventory sheets.

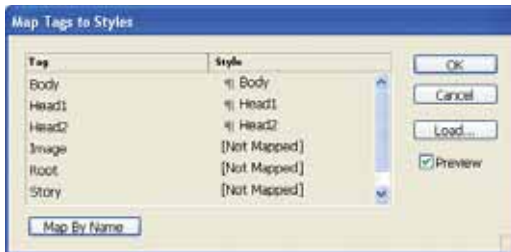
Do not import contents of whitespace-only elements—Leaves any existing content in place if the matching XML content contains only whitespace (such as a return or tab character). Use this option if you've included text between elements that you want preserved. For example, when laying out recipes generated from a database, you might add labels, such as "Ingredients" and "Instructions." As long as the parent element that wraps each recipe contains only whitespace, InDesign leaves the label in place.

Delete elements, frames, and content that do not match imported XML—Removes elements that don't have matches in the XML file from the Structure pane and from the layout. For example, when importing names and addresses, you might have placeholder text for an element containing the company name. If one of the names doesn't include the company element, InDesign deletes the element containing the placeholder text.

Mapping tags to style

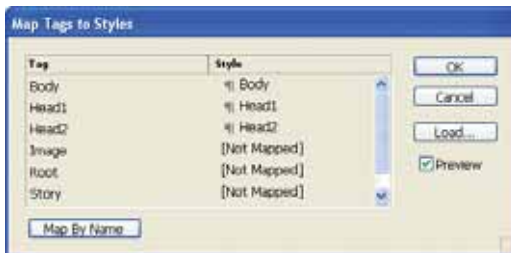
Now that you have imported XML into the document, you will format and style the text. Because the document is structured with tags, you will automate the process of formatting by establishing a relationship between the tags and styles.

1 In the Structure pane, click the menu option in the upper right corner and choose Map Tags to Styles. The Map Tags to Styles window opens.



2 In the Map Tags to Styles window, click the Map By Name button. Because this document already includes paragraph styles with names that are identical to the XML tags, InDesign CS2 can apply styles to this text. Keep the Map Tags to Styles window open.

3 Click the Preview checkbox. The tagged text becomes formatted as the Paragraph styles are applied to it. Click OK.



4 Choose File > Save.

Using XML snippets

A snippet is an XML file that is a full representation of InDesign content, including page items and any XML structure applied to those page items. A snippet lets you reuse the content, formatting, tags, and structure of a document. You can store snippets in an Object library, and place them in other documents. Use snippets to easily reuse objects—even those with complex formatting.

- 1 Choose Window > ski_brochure.indd to switch to the original document you worked on earlier in this lesson.
- 2 Choose the Selection tool (⬇️), and click anywhere on the document.
- 3 Choose View > Fit Page in Window.
- 4 Arrange your application window (Windows) or document window (Mac OS) so that you can see both your desktop and the InDesign layout.
- 5 Click to select the outline image of the snowboarder to the left of the headline. The outline of the skier and the arc across the top of the page all become selected, as they are part of a group.
- 6 Click and drag the two images and arc to the desktop of your computer. An icon appears on the desktop with a name that begins as Snippet. This XML snippet file includes all the information necessary to reproduce these images in another InDesign layout.



- 7 Choose Window > online_brochure.indd. If necessary, arrange your InDesign windows so that you can see both the snippet file you dragged to your desktop as well as the InDesign layout.

- 8 Click and drag the snippet file from your desktop onto your layout. The images and the arc are imported into the document.

Snippets retain the x and y coordinates of the original objects, making them perfect for objects that need to be reused in multiple versions of an identical document. In this case, we are using a different sized layout, so we'll move the images.



- 9 Click and drag the outline image of the snowboarder and position it to the left of the Snow & Ski Report headline text.

- 10 Choose File > Save.

Congratulations. You have completed this lesson.

Exploring on your own

- 1 Explore the Structure pane by right-clicking (Windows) or Ctrl+clicking (Mac OS) the various tags. Examine the options available in the Structure pane menu.
- 2 Apply XML tags to your own document and export the XML. Practice importing the XML into another InDesign layout.
- 3 Import XML into an InDesign layout and map the tags to InDesign styles used in the document, formatting the XML text.

Review

▶ Review questions

- 1** How and why are XML tags used in InDesign CS2?
- 2** What steps are necessary before an InDesign CS2 layout can be exported as XML?
- 3** What is the Structure pane, and what role does it play in exporting or importing XML?

▶ Review answers

- 1** XML tags are applied to content in an InDesign CS2 layout, including text, graphics and frames. Tags describe the content, allowing for it to be exported as XML. Additionally, tagged frames can have XML imported into them.
- 2** Prior to exporting an InDesign CS2 layout as XML, all objects that will be a part of the XML must be tagged. Additionally, the sequence of the XML tags must be verified in the Structure pane, and any attributes necessary must be added in the Structure pane.
- 3** The Structure pane provides a listing of all tagged objects used in an InDesign CS2 layout. It is used to import, manage and export XML. It is used to organize the hierarchy of tagged objects, apply attributes, and establish the relationship between tags and styles.

© 2006 Adobe Systems Incorporated and its licensors. All rights reserved.

Adobe® InDesign® CS2 Classroom in a Book® for Windows® and Mac OS

If this guide is distributed with software that includes an end user agreement, this guide, as well as the software described in it, is furnished under license and may be used or copied only in accordance with the terms of such license. Except as permitted by any such license, no part of this guide may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, recording, or otherwise, without the prior written permission of Adobe Systems Incorporated. Please note that the content in this guide is protected under copyright law even if it is not distributed with software that includes an end user license agreement.

The content of this guide is furnished for informational use only, is subject to change without notice, and should not be construed as a commitment by Adobe Systems Incorporated. Adobe Systems Incorporated assumes no responsibility or liability for any errors or inaccuracies that may appear in the informational content contained in this guide.

Please remember that existing artwork or images that you may want to include in your project may be protected under copyright law. The unauthorized incorporation of such material into your new work could be a violation of the rights of the copyright owner. Please be sure to obtain any permission required from the copyright owner.

Any references to company names in sample files are for demonstration purposes only and are not intended to refer to any actual organization.

Adobe, the Adobe logo, Acrobat, Classroom in a Book, Illustrator, InDesign, Myriad, PageMaker, Photoshop, PostScript and PostScript 3 are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.

Apple, Mac, and Macintosh are trademarks of Apple Computer, Inc. registered in the U.S. and other countries. Microsoft, Windows, and Windows NT are trademarks of Microsoft Corporation registered in the U.S. and/or other countries. TRUMATCH is a registered trademark of Trumatch, Inc. All other trademarks are the property of their respective owners.

Adobe Systems Incorporated, 345 Park Avenue, San Jose, California 95110-2704, USA

Notice to U.S. Government End Users. The Software and Documentation are "Commercial Items," as that term is defined at 48 C.F.R. §2.101, consisting of "Commercial Computer Software" and "Commercial Computer Software Documentation," as such terms are used in 48 C.F.R. §12.212 or 48 C.F.R. §227.7202, as applicable. Consistent with 48 C.F.R. §12.212 or 48 C.F.R. §§227.7202-1 through 227.7202-4, as applicable, the Commercial Computer Software and Commercial Computer Software Documentation are being licensed to U.S. Government end users (a) only as Commercial Items and (b) with only those rights as are granted to all other end users pursuant to the terms and conditions herein. Unpublished-rights reserved under the copyright laws of the United States. Adobe Systems Incorporated, 345 Park Avenue, San Jose, CA 95110-2704, USA. For U.S. Government End Users, Adobe agrees to comply with all applicable equal opportunity laws including, if appropriate, the provisions of Executive Order 11246, as amended, Section 402 of the Vietnam Era Veterans Readjustment Assistance Act of 1974 (38 USC 4212), and Section 503 of the Rehabilitation Act of 1973, as amended, and the regulations at 41 CFR Parts 60-1 through 60-60, 60-250, and 60-741. The affirmative action clause and regulations contained in the preceding sentence shall be incorporated by reference.

Adobe Press books are published by Peachpit, Berkeley, CA. To report errors, please send a note to errata@peachpit.com.