

Writing Accessible PDFs Using Word

Authored by
Helmut Hoss, Kim Andersen & Catherine Bridge

1st ed. July 2020

DOI: 10.26288/5f9a212740356

ISBN: 978-0-7334-3947-6



www.homemods.info

Abstract

This guide to writing accessible PDFs in Microsoft Word is provided to give users the tools to make their documents accessible to a wider audience. It is intended for anyone who creates PDF documents for public use, using Microsoft Word as the document creator and who is familiar with that software. While the focus is on Microsoft Word the principles outlined in the guide may be applied to other document creation software at the user's discretion.

The use of PDF remediation software to finalise the document and make it fully compliant with the relevant standards is recommended.

Keywords

PDF Accessibility; Document Creation; Standards; Microsoft Word; Adobe

Contribution of Authors

This is the first edition of the Occasional Paper: Writing Accessible PDFs Using Word for the Home Modification Information Clearinghouse, UNSW Sydney.

Helmut Hoss wrote the guide and tested the advice contained therein.

Kim Andersen edited the guide.

Catherine Bridge supervised the project and edited the guide.

Acknowledgements

This material has been published by the Home Modification Information Clearinghouse within the faculty of Built Environment, UNSW Sydney (University of New South Wales).

This research was supported fully by the Australian Government through the Australian Research Council's Linkage Infrastructure, Equipment and Facilities Projects funding scheme (project LE180100094). The views expressed herein are those of the authors and are not necessarily those of the Australian Government or Australian Research Council.

Liability Statement

The Home Modification Information Clearinghouse team gives no warranty that the information or data supplied contain no errors. However, all care and diligence has been used in processing, analysing and extracting the information. The Home Modification Information Clearinghouse will not be liable for any loss or damage suffered upon the direct or indirect use of the information supplied in this document.

Reproduction of Material

This material has been published by the Home Modification Information Clearinghouse within the faculty of Built Environment, UNSW Sydney (University of New South Wales).

Any table or material published in this Occasional Paper may be reproduced and published without further license, provided that due acknowledgement is made of this source. The preferred acknowledgment style is:

Hoss, H., Andersen, K. & Bridge, C. (2020) Writing Accessible PDFs Using Word 1 ed. Sydney: Home Modification Information Clearinghouse, UNSW Sydney. (July) [online]. DOI: 10.26288/5f9a212740356 Available from HMinfo website: www.homemods.info

Contents

Abstract	1
Keywords	1
Contribution of Authors	2
Acknowledgements	2
Liability Statement	3
Reproduction of Material	3
Contents	4
Figures	5
Tables 5	
1. Introduction	6
1.1. Disclaimer	7
1.2. Software Requirements	8
2. Accessible Word Document Guidelines	8
2.1. Preparing Styles	9
2.2. Applying Styles	12
2.3. Preparing Document Structure	15
2.4. Adding Content	18
2.5. Inspecting Document	29
3. Generating Accessible PDFs from Word Documents	32
3.1. Word to PDF	32
3.2. Inspect PDF Accessibility Issues	33
3.3. Repair PDF Accessibility Issues	35
Appendix	55
Appendix A	55
Appendix B	55
Appendix C	56

Figures

Figure 1. Overall process for achieving PDF/UA compliance	7
Figure 2. How to prepare your document's styles	10
Figure 3. How to prepare your document's structure	15
Figure 4. How to add content to your document	18
Figure 5. How to inspect your Word document for accessibility issues	30
Figure 6. How to convert your Word document to PDF	32
Figure 7. How to inspect your PDF for accessibility issues	34
Figure 8. Overall process for repairing PDF accessibility issues	37
Figure 9. How to repair PDF metadata and setting errors	39
Figure 10. How to repair PDF role mapping errors	43
Figure 11. How to repair PDF logical structure errors	47

Tables

Table 1. Common substitutions	14
Table 2. Example languages and language tags	42
Table 3. Example role mappings	44
Table 4. Bounding box value characteristics	49
Table 5. Common standard structure tags and their default placement layout attribute values	50

1.Introduction

Microsoft Word 2013, 2016 and Office 365 (Microsoft Word 2019) are able to output documents in PDF format, however without prior preparation and further processing, the PDF documents generated by Word will be unable to conform to PDF/Universal Accessibility standards (ISO 14289-1:2014).

Document authors must complete 3 major steps to reach a successful outcome:

- Write Word documents using appropriate technique.
- Output a partially compliant PDF via Word.
- Edit the partially compliant PDF using a PDF remediation tool to finalise the PDF.

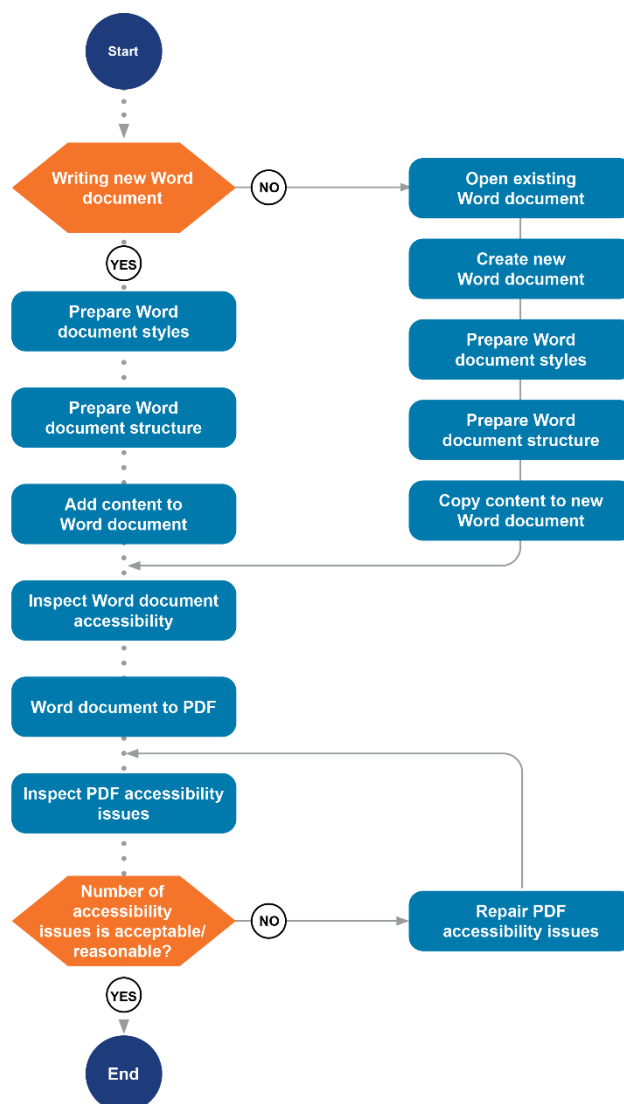


Figure 1. Overall process for achieving PDF/UA compliance

This guide assumes that you have previous experience writing documents using Microsoft Word. It is not intended to be a Microsoft Word training manual.

1.1. Disclaimer

The instructions found in this guide were developed and tested using Microsoft Windows operating systems using Australian region and language settings. While it may be possible to follow the instructions on other operating systems such as macOS or using languages other than Australian English, the generated PDFs may not be accessible.

1.2. Software Requirements

Instructions in this guide require the user to have the following software installed on their workstation:

- Microsoft Word 2013 or Microsoft Word 2016 or Microsoft Office 365 (Microsoft Word 2019)
- PDF Accessibility Checker 3 (PAC 3) version 3.0.7.0 or higher. Free download available from the Access for All website: <https://www.access-for-all.ch/en/>
- Colour Contrast Analyser version 2.2a or higher. Free download available from the Vision Australia website: <https://www.visionaustralia.org>
- Software capable of “Tagged PDF” remediation. More information available in Repair PDF Accessibility Issues.

2. Accessible Word Document Guidelines

This section describes techniques that should be employed when writing Word documents. Following these guidelines will improve the accessibility of your Word documents and facilitate the overall process of generating accessible PDFs.

You will likely be using your organisation’s branding when writing Word documents by either using a Word template or following branding steps supplied in an internal document.

It is recommended that you initially complete the instructions in this section without using your organisation’s Word templates as this will allow you to familiarise yourself with methods for achieving accessible Word documents – a prerequisite to generating accessible PDFs. Once you are comfortable generating accessible Word documents, consider adjusting or re-creating your organisation’s Word templates based on your experience following this guide.

2.1. Preparing Styles

There are over one hundred styles in Word, many which are not used in simple documents. Common styles include strong text, emphasised text, titles and headings.

Consider the different colours, font sizes and other design elements that you will be using in your document. For instance, you may be intending to use a different colour for each heading level.

Your choice of colours, font sizes and other design elements should take into account the following guidance:

- Keep formatting consistent and minimal. Too many variations in fonts, emphasis styles and indentation can distract audiences from the content.
- Do not use colour as the sole method for denoting meaning. Audiences with vision impairments may miss differences in colours. It is preferable to use variations in font sizes or styles such as small caps, italics or bold.
- Underlining should be used to denote linked text. Do not underline text for aesthetic reasons. Audiences reading electronic versions of documents may mistake the text for linked text.
- Only use one of the following text justifications: “Align Left”, “Align Right” or “Align Center”. Text that is “Justified” (evenly distributed between margins) may be difficult to read, especially by people with dyslexia.

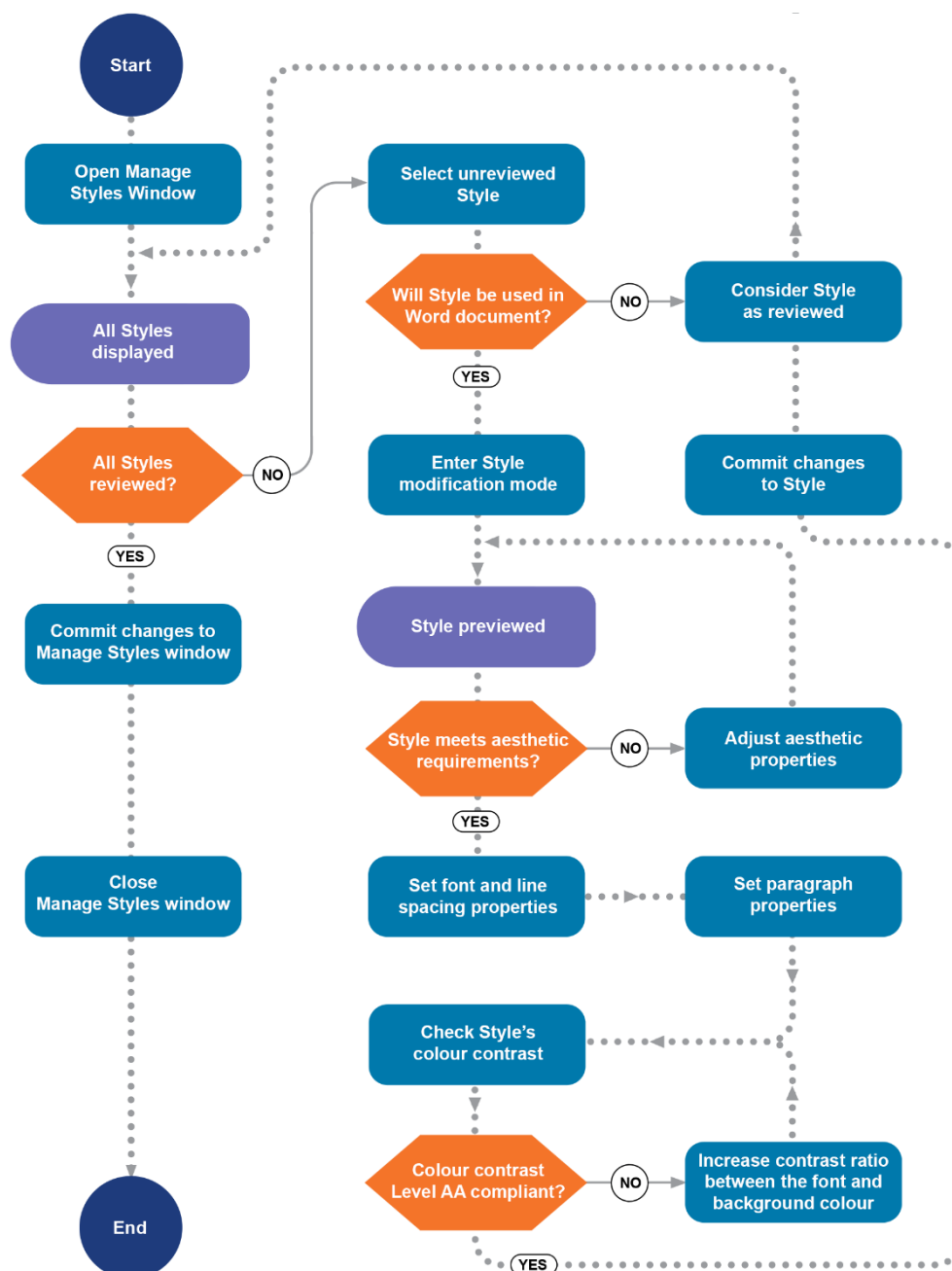


Figure 2. How to prepare your document's styles

While being mindful of the expected content in your document and the aforementioned guidance, choose and adjust styles that you foresee using by following these steps:

1. Click on the Home tab in the ribbon.
2. Under the Styles group, click on the expand icon. The Styles pane will be displayed.

3. Click on the Manage Styles button in the Styles pane.
4. The Manage Styles window will pop-up.
5. Scroll through the list of styles and note the styles that you may use. e.g. Title, Subtitle, Heading 1, Heading 2, Emphasis, Strong.
Note: The “Normal” style defines the visual appearance for all text that has not been assigned a style. All styles inherit settings from the Normal style.
6. For each style that you intend to use:
 - a. Click on the style in the list of styles.
 - b. Click on the Modify button.
 - c. Use the preview area to review the aesthetics of the style.
 - d. If the style meets your aesthetic requirements, go to the next step, otherwise make adjustments to the style before continuing.
 - e. Ensure that the following values are configured for the style:
 - Font size above 7.5pt.
 - Line spacing at least 1.5 times the font size. e.g. exactly 16.5pt line spacing for 11pt text.
 - Paragraph spacing should equal to or exceed the line spacing value. e.g. 18pt paragraph spacing minimum when the line spacing is 18pt.
 - Font character spacing at least 0.12 times the font size using expanded spacing mode. e.g. 1.45pt character spacing for 12pt text.

- Left, Centre or Right alignment – do not align text using Justify.
- f. Launch the Colour Contrast Analyser. Refer to Software Requirements for more information.
 - g. Use the Colour Contrast Analyser to check the preview of the style for contrast issues.
 - h. If contrast issues are present:
 - Change the Font Colour of the style with the aim of increasing the contrast ratio between the font colour and the background colour. e.g. If the font colour is red on a white background, attempt to use darker reds.
 - Re-check the style using the Colour Contrast Analyser and repeat these steps until no contrast issues are detected.
 - i. Click on OK to partially commit the changes.
7. Click on OK to close the Manage Styles window and commit all style changes.

Since many default styles provided by Word are lacking colour contrast testing, you will most likely encounter colour contrast issues when following the style preparation process.

2.2. Applying Styles

Conventionally, font, colour and size changes to text would be performed by using buttons located in the Font group in the Home tab in the ribbon. While it remains possible to use the conventional font controls, using them will increase the complexity of maintaining aesthetic consistency throughout the document and potentially reduce accessibility.

Styles are able to define the structure of a document while conveying emphasis of content to screen readers, therefore it is imperative to substitute the use of font controls with the use of styles when changing the visual appearance of text.

The default visual appearance of text in a Word document is defined by the “Normal” style. All text that is entered into a document without any style being preselected will be automatically styled as “Normal”.

To change the style of text:

- Click on the Home tab in the ribbon.
- Highlight a range of text or click on a sentence that you wish to change.
- Select a style from the Styles group in the Home tab.
If the style you would like to use is not displayed:
 - Click on the “More” drop-down arrow button in the Styles group.
 - Click on Apply Styles. The Apply Styles pane will display.
 - Select your preferred style from the Style Name drop-down list.

Example Substitutions

Word provides predefined styles for commonly used formats consisting of text size and colour combinations.

Table 1. Common substitutions

Intention	Conventional Method	Preferred Method
Bold text	Select text and click on the Bold button.	Change text to Strong style.
Italicise text	Select text and click on the Italic button.	Change text to Emphasis style.
Add emphasis to italicised text	Select text, click on the italic button and then choose a colour from Font Colour button.	Change text to Intense Emphasis style.
Add a heading	Select text, click on the Bold button and then click on the Increase Font Size button.	Change text to a Heading style that corresponds with the heading depth. e.g. Select Heading 1 for top level headings.

2.3. Preparing Document Structure

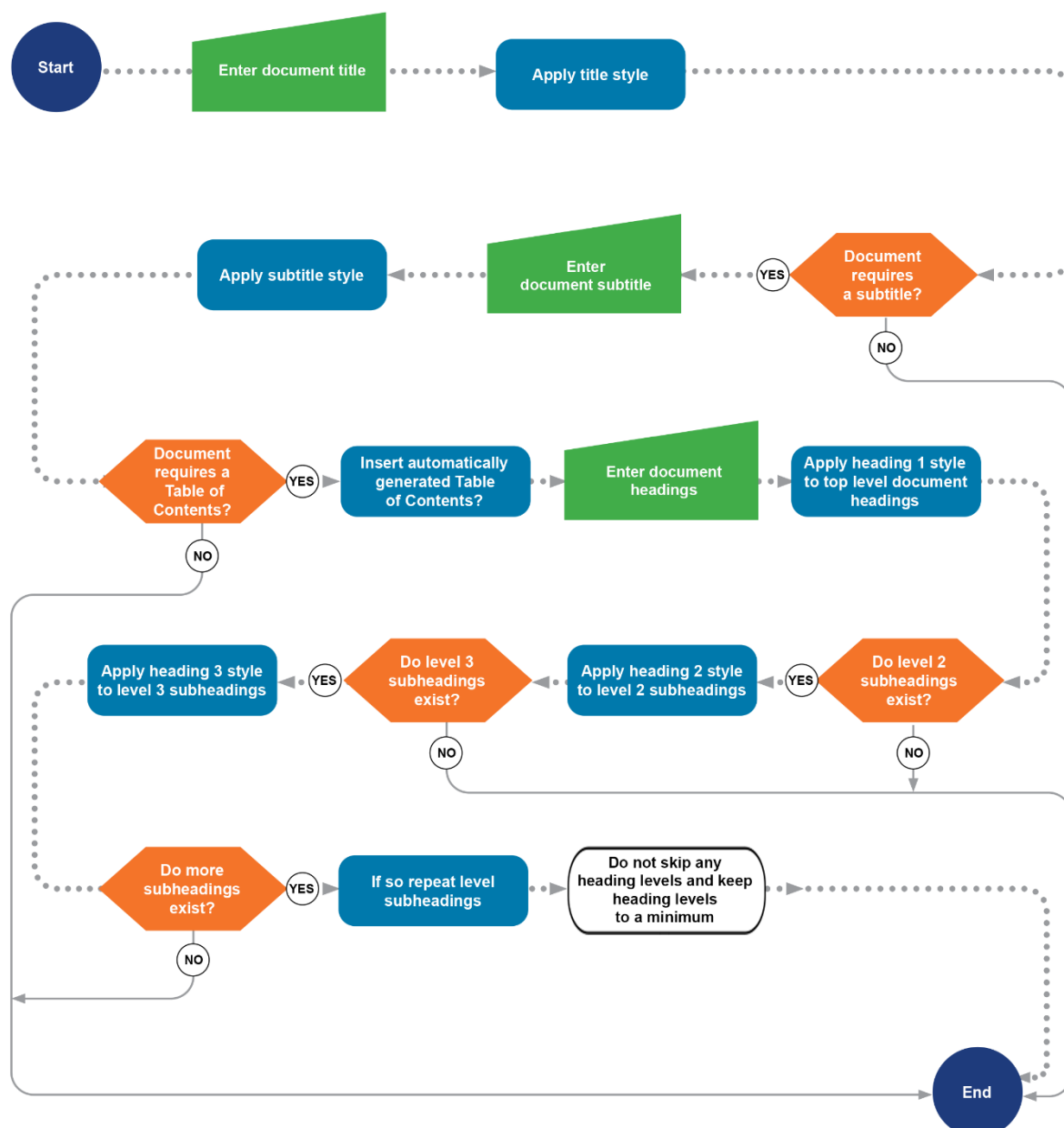


Figure 3. How to prepare your document's structure

Once you have prepared the styles that you will be using in your document, the next step is to prepare your document structure.

Establishing the structure of your document before writing content is often an efficient method of tracking your progress.

A typical document structure may consist of the following elements in the order listed below:

- Title
- Subtitle
- Table of Contents
- Headings
 - Subheadings

2.3.1. Title and Subtitle

Title and subtitle styles are listed as recommended styles in Word, however they currently do not offer any additional structure information to Word documents or PDFs. Using title and subtitle styles may be beneficial in future versions of Word and thus should be used when defining the main and secondary titles in a document.

2.3.2. Table of Contents

Word is able to automatically generate a table of contents that includes internal links to headings found within your document. Automatically generated internal links are preserved during Word to PDF conversion.

2.3.3. Headings and Subheadings

Word allows for up to 9 heading styles to be defined, starting with Heading 1 and ending with Heading 9. Heading 1 and Heading 2 styles are listed as recommended styles in Word.

Heading 1 is the highest/top-most heading level.

Heading 9 is the lowest/bottom-most heading level.

It is recommended to use no more than 4 heading levels in your document in order to help your audience track the context of each section easily as they read and navigate your document. If your document exceeds 3 heading levels, review your content and if possible trim and combine content to eliminate excess heading levels.

Use heading styles correctly to ensure that document structure is preserved during Word to PDF conversion and is ultimately understood by screen readers:

- Do not use heading styles to emphasise text. For example, do not use the Heading 1 style to bold and enlarge an important sentence, use either the Strong style or the Emphasis style.
- Do not skip heading levels. First-level headings should be styled using the Heading 1 style and should only contain sub-headings styled using the Heading 2 style. Second-level headings should be styled using the Heading 2 style and should only contain sub-headings styled using the Heading 3 style.
- Do not use Heading 7 to Heading 9 since PDF formats do not recognise heading levels beyond Heading 6 and therefore any Word to PDF conversion process will fail to map Heading 7 to Heading 9 directly.

2.4. Adding Content

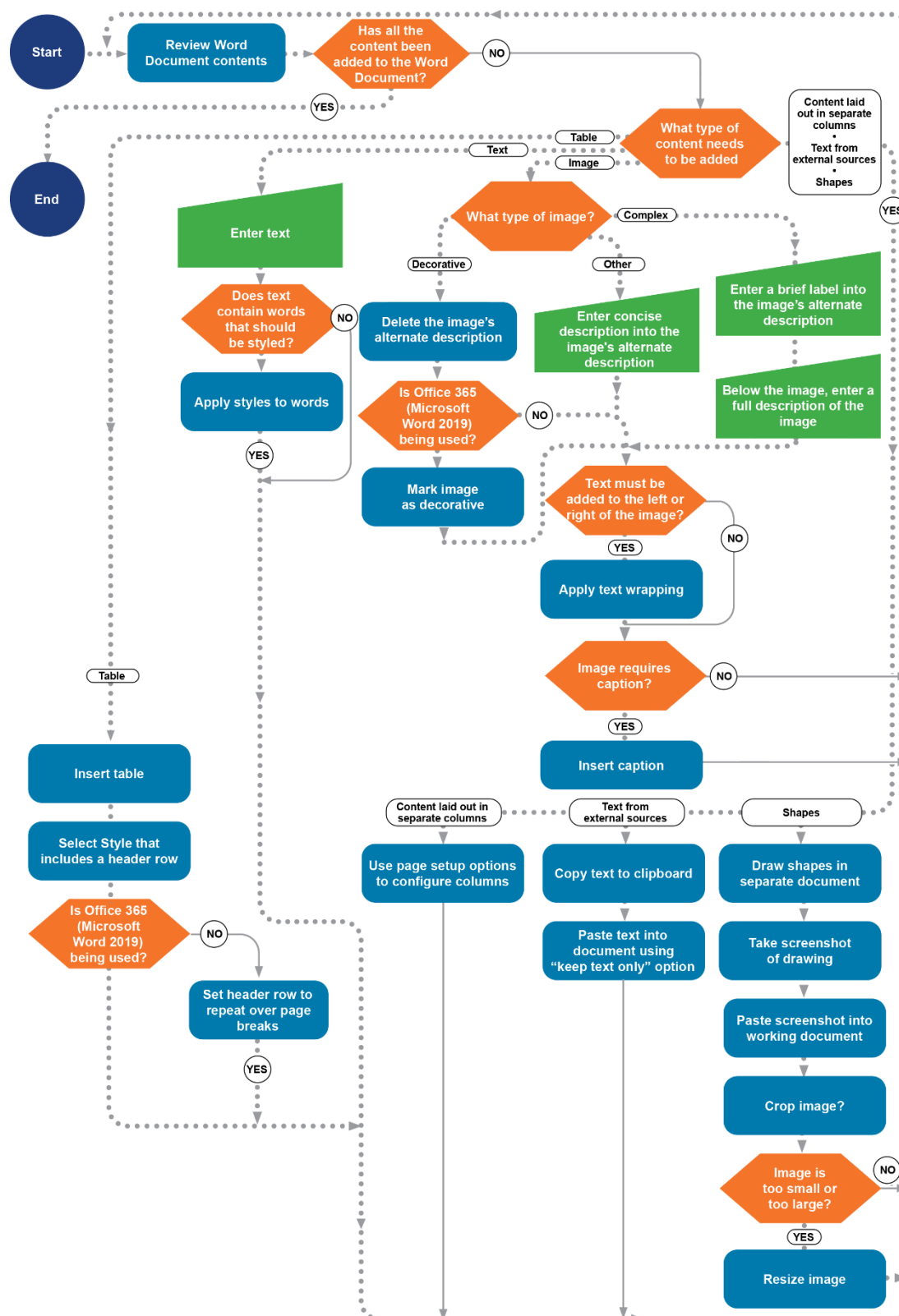


Figure 4. How to add content to your document

Text in your Word document is likely to mostly be styled using the “Normal” style. Additional care must be taken when adding other types of content or structures to your document.

2.4.1. Images

Before you add an image to your document, note the following:

- Images should not be added inside tables that only serve to position them or inside single celled tables created to produce caption boxes - tables should not be used as a layout tool.
- Do not use images of text to substitute text content.
- Images that have text wrapping enabled will produce extra accessibility issues to fix. If the issues are not fixed, some screen readers will be unable to determine the correct reading order of the image and wrapped text. You may choose to reduce the complexity of your document by not using images wrapped with text.

After adding an image to your document, follow the steps below depending on your version of Microsoft Word/Office.

Microsoft Word 2013

1. Right-click on the image.
2. Click on “Format Picture...” in the pop-up context menu.
3. In the Format Picture panel, click on the Layout & Properties icon (third icon from the left).
4. Click on Alt Text.
5. Blank the Title field as it is no longer used and could confuse people who may edit your document in the future. The Title field was used to provide a preview of the information contained in the

Description field. Word to PDF conversion software omits information in the Title field and thus it is helpful to promote the Description field as the true source of alternative text information.

6. If the image is decorative, such as a logo or separator:
 - a. Blank the Description field as this will indicate to screen readers that the image may be decorative.
7. If the image is complex, such as a chart or technical diagram:
 - a. Enter a brief label for the image into the Description field.
e.g. Anterior View of Skeletal System
 - b. Briefly describe the details of the image using text. Add the text under the image. If the image is an “info graphic”, provide the information it contains in text format – do not rely on the image being visible to all audiences.
8. If the image is none of the above, such as a product image or a photo of a person:
 - a. Enter a description of the image into the Description field.
The description should be concise and under 100 characters in length.
9. If you would like to wrap text on the left or the right of the image:
 - a. Click on the image.
 - b. Click on the Format tab in the ribbon.
 - c. Inside the Arrange group, click on Wrap Text and then “More Layout Options...”. The Layout dialog will be displayed.
 - d. Click on the Text Wrapping tab.

- e. Click on Square under the Wrapping style heading.
 - f. Under the Wrap text heading:
 - Click on “Both sides” if the image will have text placed on either side.
 - Click on “Left only” if the image will have text placed on its left hand side.
 - Click on “Right only” if the image will have text placed on its right hand side.
 - g. Modify the values under the “Distance from text” heading to ensure that text is not joined to the image. e.g. Set the Right value to 0.3 cm if the image will have text placed on its right hand side.
 - h. Click on the OK button.
10. If you would like to add a caption to the image:
- a. Right-click on the image.
 - b. Click on “Insert Caption...” in the pop-up context menu.
 - c. Enter the caption text into the Caption field. e.g. Figure 1.
Flowchart
 - d. Click on the OK button.

Microsoft Word 2016

1. Right-click on the image.
2. Click on “Picture...” in the pop-up context menu.
3. Click on the Alt Text tab.
4. If the image is decorative, such as a logo or separator:
 - a. Blank the Alternative Text field as this will indicate to screen readers that the image may be decorative.
 - b. Click on the OK button to close the Format Picture dialog.
5. If the image is complex, such as a chart or technical diagram:
 - a. Enter a brief label for the image into the Alternative Text field. e.g. Anterior View of Skeletal System
 - b. Click on the OK button to close the Format Picture dialog.
 - c. Briefly describe the details of the image using text. Add the text under the image. If the image is an “info graphic”, provide the information it contains in text format – do not rely on the image being visible to all audiences.
6. If the image is none of the above, such as a product image or a photo of a person:
 - a. Enter a description of the image into the Alternative Text field. The description should be concise and under 100 characters in length.
 - b. Click on the OK button to close the Format Picture dialog.

7. If you would like to wrap text on the left or the right of the image:

- a. Click on the image.
- b. Click on the Format tab in the ribbon.
- c. Inside the Arrange group, click on Wrap Text and then “More Layout Options...”. The Layout dialog will be displayed.
- d. Click on the Text Wrapping tab.
- e. Click on Square under the Wrapping style heading.
- f. Under the Wrap text heading:
 - Click on “Both sides” if the image will have text placed on either side.
 - Click on “Left only” if the image will have text placed on its left hand side.
 - Click on “Right only” if the image will have text placed on its right hand side.
- g. Modify the values under the “Distance from text” heading to ensure that text is not joined to the image. e.g. Set the Right value to 0.3 cm if the image will have text placed on its right hand side.
- h. Click on the OK button.

8. If you would like to add a caption to the image:

- a. Right-click on the image.
- b. Click on “Insert Caption...” in the pop-up context menu.

- c. Enter the caption text into the Caption field. e.g. Figure 1.
Flowchart
- d. Click on the OK button.

Microsoft Office 365 (Microsoft Word 2019)

1. Right-click on the image.
2. Click on “Edit Alt Text...” in the pop-up context menu.
3. If the image is decorative, such as a logo or separator:
 - a. Tick the “Mark as decorative” check box.
4. If the image is complex, such as a chart or technical diagram:
 - a. Enter a brief label for the image into the Alt Text field. e.g.
Anterior View of Skeletal System
 - b. Briefly describe the details of the image using text. Add the text under the image. If the image is an “info graphic”, provide the information it contains in text format – do not rely on the image being visible to all audiences.
5. If the image is none of the above, such as a product image or a photo of a person:
 - a. Enter a description of the image into the Alt Text field. The description should be concise and under 100 characters in length.
6. If you would like to wrap text on the left or the right of the image:
 - a. Click on the image.

- b. Click on the Format tab in the ribbon.
 - c. Inside the Arrange group, click on Wrap Text and then “More Layout Options...”. The Layout dialog will be displayed.
 - d. Click on the Text Wrapping tab.
 - e. Click on Square under the Wrapping style heading.
 - f. Under the Wrap text heading:
 - Click on “Both sides” if the image will have text placed on either side.
 - Click on “Left only” if the image will have text placed on its left hand side.
 - Click on “Right only” if the image will have text placed on its right hand side.
 - g. Modify the values under the “Distance from text” heading to ensure that text is not joined to the image. e.g. Set the Right value to 0.3 cm if the image will have text placed on its right hand side.
 - h. Click on the OK button.
7. If you would like to add a caption to the image:
- a. Right-click on the image.
 - b. Click on “Insert Caption...” in the pop-up context menu.
 - c. Enter the caption text into the Caption field. e.g. Figure 1. Flowchart

- d. Click on the OK button.

2.4.2. Tables

Tables should only be used to provide information in tabular format, they should not be used for layout purposes such as creating page columns, captioned boxes and forms since Word provides other methods to achieve a variety of layouts.

It is imperative for tables to include header cells as headers relay the context of each cell to screen readers which in turn convey the information to screen reader users as they navigate cells.

Once you have inserted a table, follow these steps:

1. Click inside the first cell of the table.
2. Click on the Design tab in the ribbon.
3. Inside the Table Style Options group, tick the Header Row check box.
4. Inside the Table Styles group, click on your preferred style.
5. If you are using Microsoft Office 2013 or Microsoft Office 2016:
 - a. Select the entire header row in your table.
 - b. Click on the Layout tab in the ribbon.
 - c. Click on the Properties button inside the Table group. The Table Properties dialog will pop-up.
 - d. Click on the Row tab.

- e. Tick the “Repeat as header row at the top of each page” check box unless it is important to not repeat the header row whenever the table is split between pages in your document.
- f. Click on the OK button.

2.4.3. Columns

If your document needs to display text in two or more columns, do not use tables to create the layout. Follow the steps below to add columns to your document:

1. Click on the Page Layout tab in the ribbon.
2. Inside the Page Setup group, click on Columns.
3. Click on the “More Columns...” menu item. The Columns dialog will display.
4. Click on a Column preset or enter values into the fields provided by the dialog to setup your preferred column configuration.
5. Click on the OK button.

2.4.4. Text from External Sources

Content that is pasted into Word via the clipboard by default will typically include the formatting present in the source, whether it is a document, webpage or spreadsheet. The included formatting will likely consist of styles that specifically pertain to the pasted content and are independent from your documents style.

Independent styles can be difficult to change and may not be responsive to changes made to your document’s style and therefore should be replaced with your document’s styles.

Furthermore, pasted content may include hidden containers, characters and even links that may cause your document to be inaccessible.

You can avoid issues caused by text from external sources by ensuring to paste only the text content into your document.

1. Copy text from an external source such as a webpage.
2. Right-click on the location in your document where you would like the copied text to be inserted.
3. Click on Paste Options | Keep Text Only in the pop-up context menu.
4. If you intend to format the text so that it resembles the source document, apply styles using the method described in this document under heading Applying Styles.

2.4.5. Shapes

While shapes offer a convenient method to draw diagrams inside documents, the shapes added to a document can limit the document's accessibility.

Ideally, diagrams consisting of various connected shapes such as flowcharts should be added to your document as a single image, followed by a text-based description.

If you are comfortable using Word to create flowcharts and other diagrams:

1. Create a new document.
2. Draw your diagrams in the new document.

3. Once you have completed your diagram, take a screenshot of it by pressing the Print Screen key on your keyboard. The screenshot will be stored in your clipboard.
4. Paste the screenshot into your document.
5. Right-click on image in your document.
6. Click on the Crop icon adjacent to the pop-up context menu. A crop selection will appear over your image.
7. Drag the corners of the crop selection until you are satisfied with the part of the image that will be shown in your document.
8. Click outside the crop selection to commit the changes.
9. If the image needs to be resized, left click and drag the bottom left hand side corner of the image into or away from the top right hand side corner of the image. Release the left click mouse button once you are satisfied with the size of the image.

2.5. Inspecting Document

If you have followed the instructions in this guide up to this point, it is likely that your Word document will be mostly accessible in practice.

Word contains a simple built-in accessibility checker that will identify issues in your document and provide links to the issues. Please note that the accessibility checker provided in Microsoft Word contains limitations. A list of limitations is available further below.

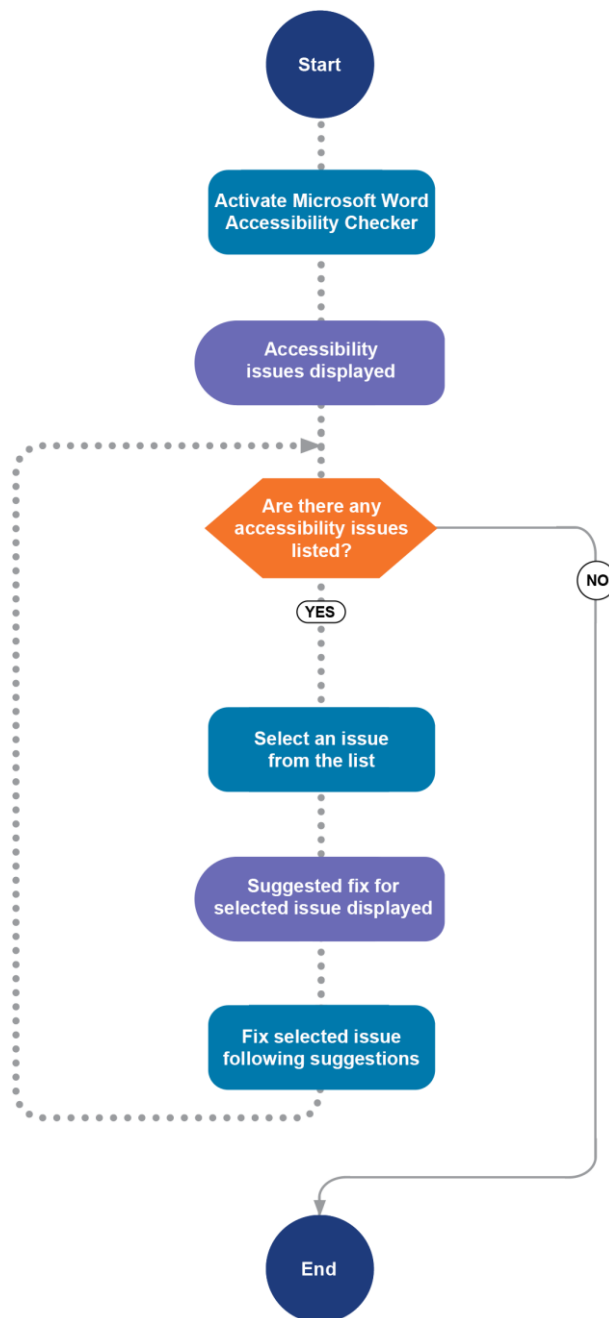


Figure 5. How to inspect your Word document for accessibility issues

To inspect your document:

1. Click on the File tab in the ribbon.
2. Click on Info.
3. Click on Check for Issues.

4. Click on Check Accessibility in the pop-up context menu.
5. Accessibility issues will be listed inside the Accessibility Checker panel located on the right hand side of your document.
6. If there are issues listed:
 - a. Click on an issue.
 - b. Additional information about the issue will be displayed at the bottom of the panel.
 - c. Action changes according to the information provided.

Microsoft Word's built-in accessibility checker will consider images that have text wrapping enabled as a potential accessibility issue. The warning/error displayed for such images will be "Object(s) not Inline" or "Image or object not inline". You may ignore these errors if you intend to continue to the next section of this guide.

2.5.1. Limitations

Microsoft Word's built-in accessibility checker has the following limitations:

- Microsoft Word 2013 and 2016 will not identify colour contrast issues.
- Microsoft Word 2013 and 2016 incorrectly consider that missing alternate text in Tables are errors.
- Microsoft Word 2013, 2016 and Office 365 (Microsoft Word 2019) will not detect heading level issues correctly if your document does not contain a heading styled as Heading 1.

3. Generating Accessible PDFs from Word Documents

The effectiveness of the instructions in this section is determined by the manner in which the Word document was written. Ideally, the Word document that you are converting to PDF was authored using the instructions found in this guide.

3.1. Word to PDF

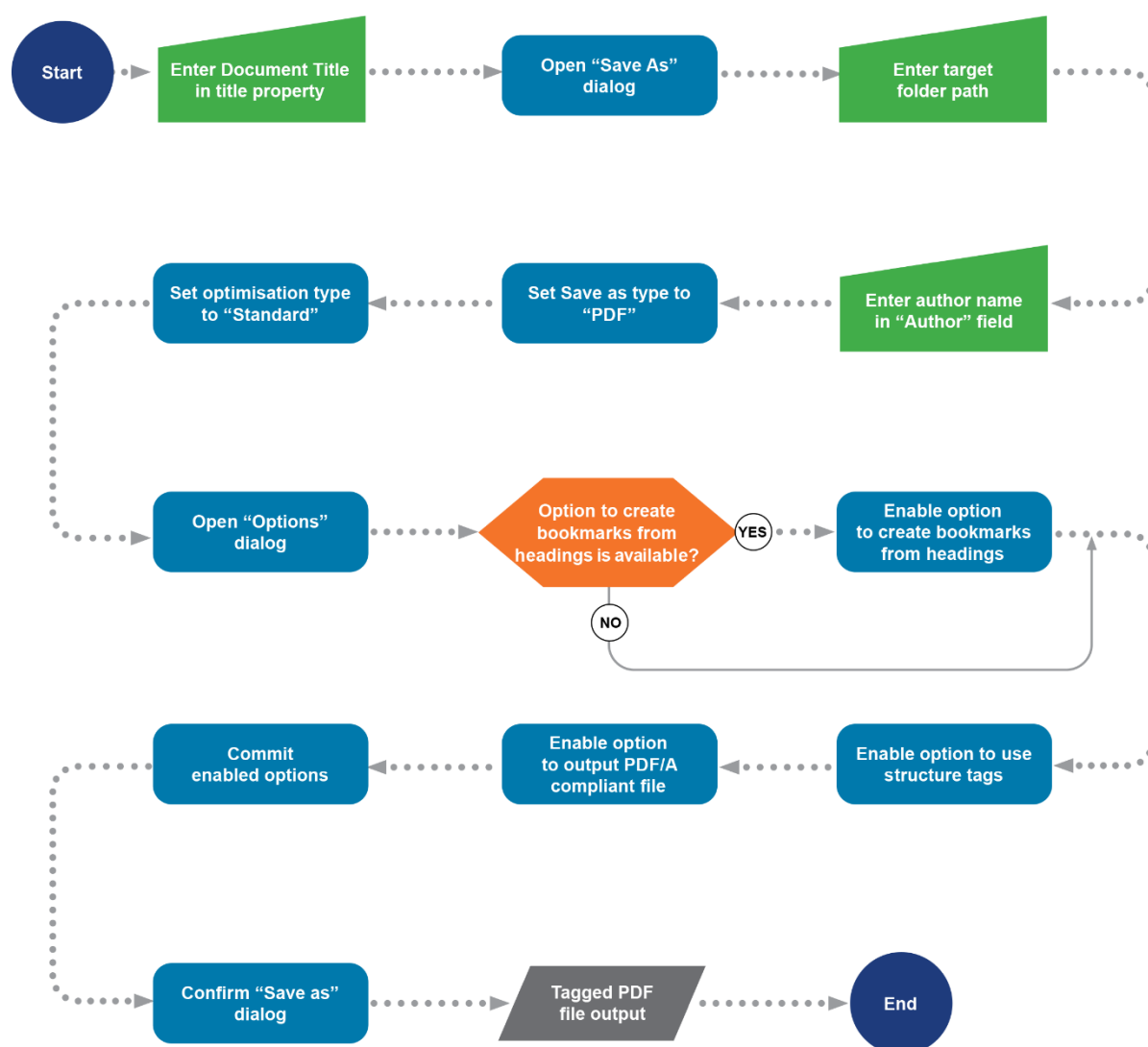


Figure 6. How to convert your Word document to PDF

Follow these steps to convert your Word document to PDF:

1. Click on the File tab in the ribbon.
2. Click on Info.
3. Under the Properties heading, check that the Title property is correct and make adjustments if necessary.
4. Click on Save As.
5. Select the location where you would like to save the output PDF.
6. Enter a file name.
7. Check the Author field. If it is incorrect, edit the field.
8. Select PDF from the Save as type drop-down list.
9. Click on the Standard option in the “Optimize for” field.
10. Click on the Options button. The Options dialog will display.
11. In the Options dialog, tick the following check boxes:
 - a. Create bookmarks using: Headings (if the option is available to select)
 - b. Document structure tags for accessibility
 - c. PDF/A compliant
12. Click on the OK button.
13. Click on the Save button.

The PDF that is output will contain invisible markers known as structure tags/structure elements that help screen readers understand the structure of your document. A PDF containing structure tags is referred to as a “Tagged PDF”.

To find out more about PDF standard structure tags that are present in Tagged PDFs refer to Appendix D.

3.2. Inspect PDF Accessibility Issues

You will need to have PDF accessibility checking software available on your workstation to complete this section of the guide. This guide is based on using “PDF Accessibility Checker” version 3.0.7.0 (PAC 3). For more information, refer to Software Requirements.

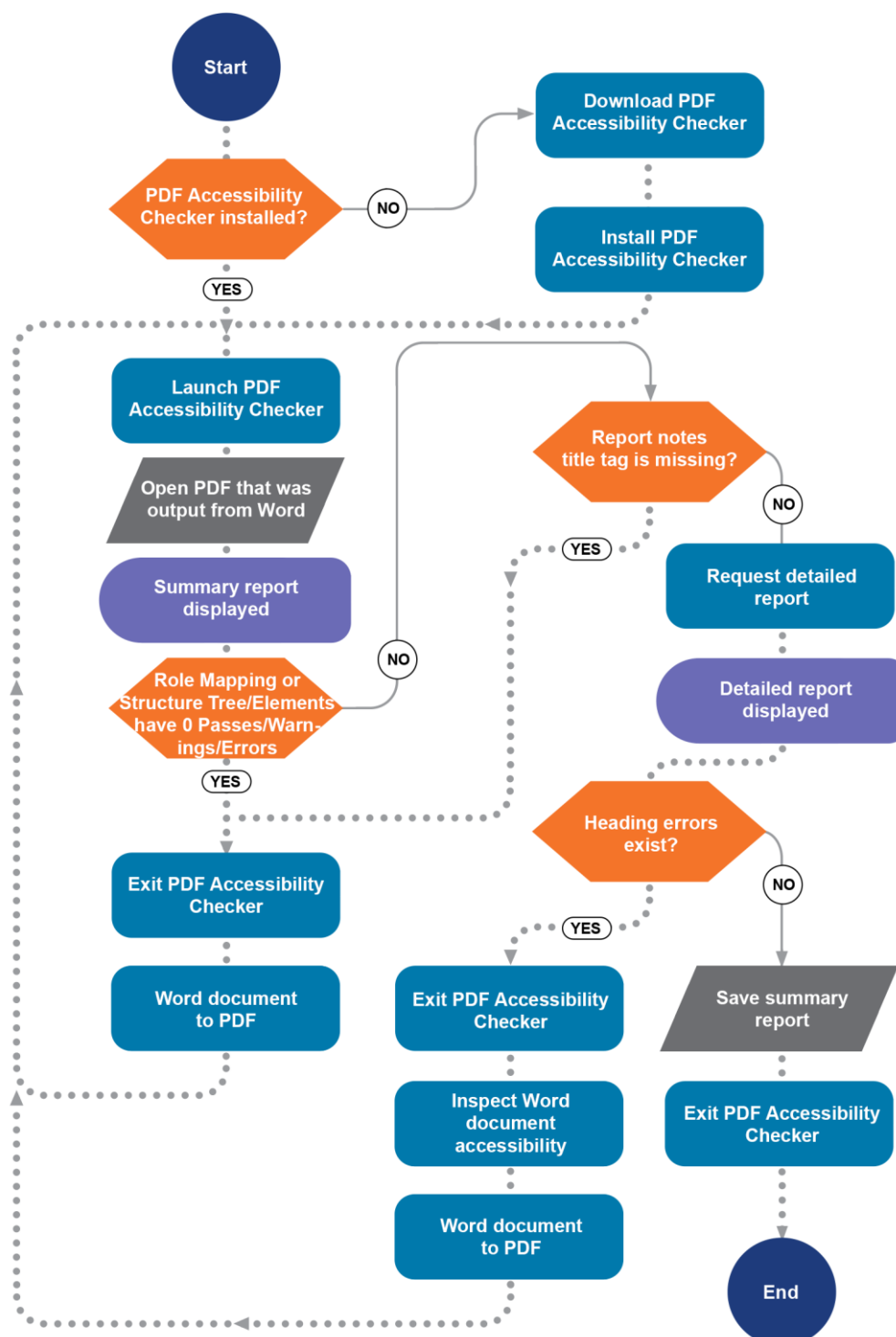


Figure 7. How to inspect your PDF for accessibility issues

Once you have downloaded and installed PDF Accessibility Checker, follow the steps below.

1. Launch PDF Accessibility Checker (PAC 3).
2. Click on the Drag and Drop icon.

3. Select the file that was output when you converted your Word document to PDF.
4. A summary report will be displayed.
5. If the summary report contains no passes, warnings and failures for Structure Tree, Role Mapping or Structure Elements, exit PAC 3 and follow this guide from Word to PDF onwards. If the issue persists, there may be a technical fault in your Microsoft Word installation.
6. If the summary report notes that there is no title, exit PAC 3 and follow this guide from Word to PDF onwards.
7. Click on the Results in Detail button.
8. If there are errors under Logical Structure | Structure Elements | Headings, exit PAC 3 and review your document's structure according to Preparing Document Structure. It is likely that your headings are not ordered correctly. Once adjustments are made, follow this guide from Inspecting Document onwards.
9. Click on the Summary Report as PDF button to save a copy of the summary report.
10. Exit PAC 3.

3.3. Repair PDF Accessibility Issues

Repairing accessibility issues in your PDF document requires PDF remediation software that is able to modify the invisible structure tags that it contains. Note that the term “structure tag” and “structure element” are interchangeable.

PDF remediation software that is fit for this purpose should:

- Allow PDF meta tags such as Title and Language tags to be edited.
- Include functions that enable editing XMP metadata, either directly or via an interface.
- Allow the mapping between non-standard structure tags and standard structure tags (known as role mapping) to be edited.
- Display the logical structure of the document based on its structure tags.
- Display a preview of the document pages and their content.
- Provide the ability to edit structure tags and their properties.
- Provide a method to edit page properties.

- Include a built-in accessibility checker that is able to analyse an open document.

Once you have sourced and installed PDF remediation software meeting the above criteria, follow the steps below under each heading to finalise your PDF.

Note that the error message text in this section of the guide is based on PAC 3's error message text. It is recommended that you save and retest the PDF after fixing an error as fixing one error sometimes also fixes related errors. For more information on PAC 3, refer to Software Requirements.

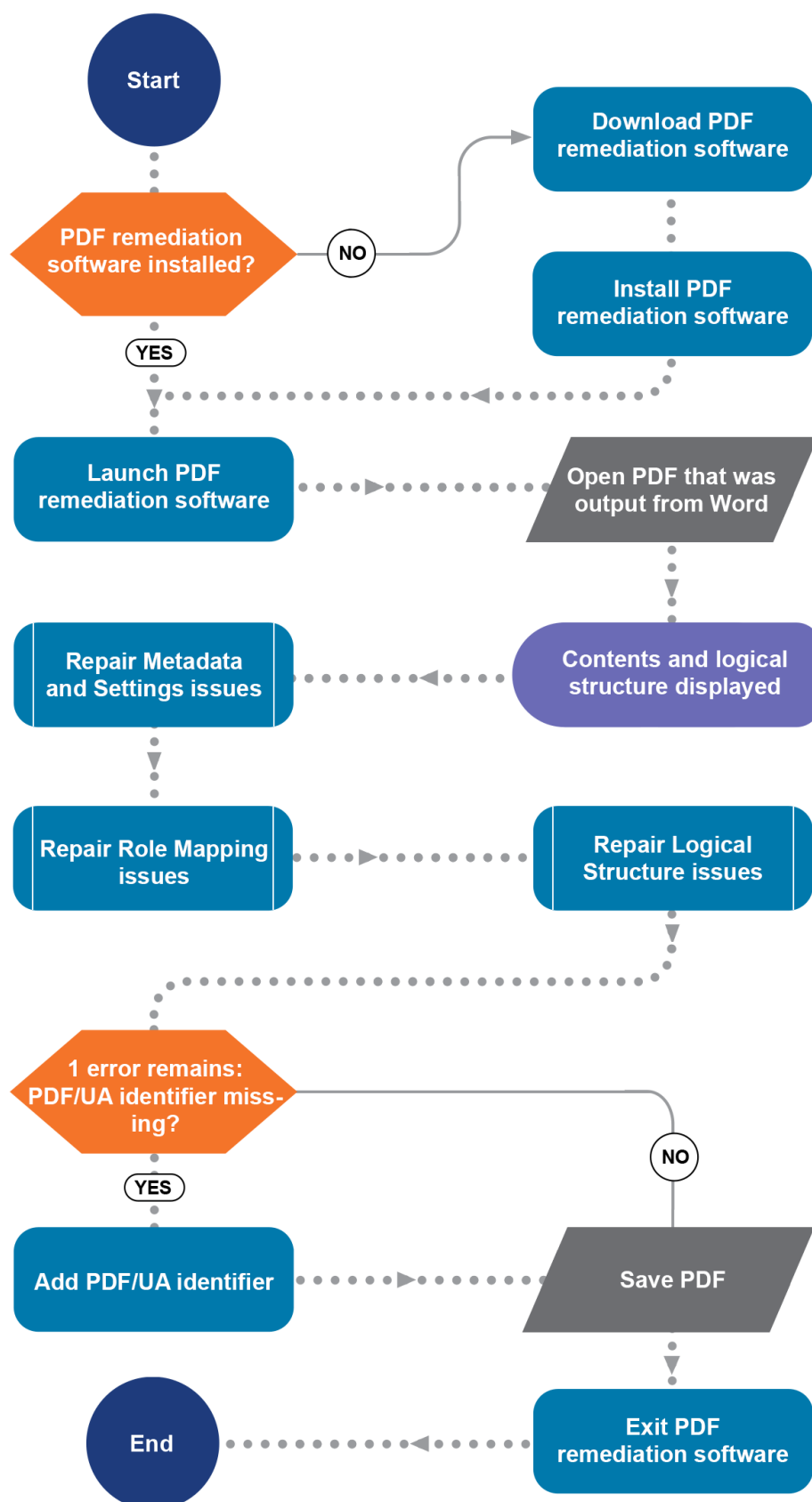


Figure 8. Overall process for repairing PDF accessibility issues

3.3.1. Initialise PDF Editing

1. Launch PDF remediation software. Refer to Repair PDF Accessibility Issues for more information on selecting adequate PDF remediation software.
2. Open the file that was output when you converted your Word document to PDF.
3. The document pages, content and logical structure will be displayed.

3.3.2. Metadata and Settings

This section describes error messages you may encounter regarding your document's metadata and setting values, and provides methods to address each error.

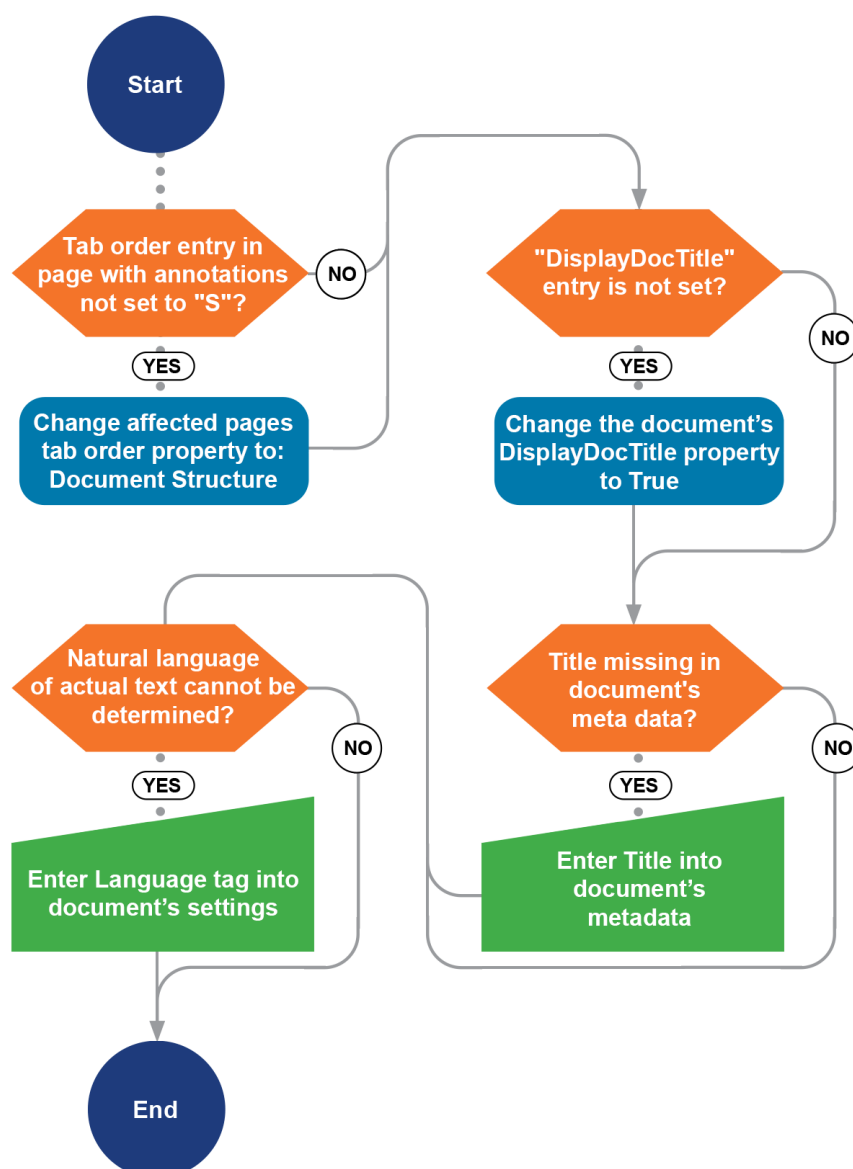


Figure 9. How to repair PDF metadata and setting errors

Tab order entry in page with annotations not set to "S" (structure)

If this error is present, check each page in your document that is marked with this error and change their tab order property to: Document Structure

Your PDF remediation software may provide a shortcut to automatically set the tab order property of all pages in your document to the correct value.

"DisplayDocTitle" entry is not set

This error is caused when document settings indicate that PDF viewers should use the document's file name in their window title. If this error is present, change the document's DisplayDocTitle property to True. A value of True indicates that the document Title tag will be displayed in the window title.

Title missing in document's XMP metadata/Title is "(no title)"

These errors are caused by the document's Title tag not being specified, therefore the solution is to provide the title of the document. If this error is present, you will need to edit the XMP metadata in your document. Your PDF remediation software may provide an interface to edit your document's Title tag via a simple form, however if no such function is provided, follow the steps below to add the Title tag manually by editing your document's XMP metadata.

Before you begin, note that these steps involve searching for text, inserting text and replacing text within XMP metadata which is a text-based format:

Look for a block of text in your document's XMP metadata that starts with <dc:title>

If you find the above text, look for the following text underneath it

```
<rdf:li xml:lang="x-default"></rdf:li>
```

If the above text exists, enter the document's title between the first greater than character and the last lesser than character. If the document title were "Windows 10 for Experts", the resulting text would resemble the following text:

```
<rdf:li xml:lang="x-default">Windows 10 for Experts</rdf:li>
```

If you were unable to locate any of the above text, look for text in the XMP metadata that starts with `<rdf:Description` and includes the text:
`xmlns:dc="http://purl.org/dc/elements/1.1/"`

Examples of how matching text may appear are available in Appendix A.

If you were able to locate a block of text matching the aforementioned criteria, append the text supplied in Appendix B immediately after the last greater than character. Otherwise, you will need to add the XMP metadata supplied in Appendix C to your document.

Ensure to replace the text “ENTER DOCUMENT TITLE HERE” in the XMP metadata with the document title.

PDF/UA identifier missing

If you encounter this error, ignore it until you reach the last section of this guide (Finalising).

Natural language of actual text cannot be determined/Language is “(no language)”

If there are a multiple errors regarding the document’s natural language or its lack of language, your document is likely missing a valid value in its Language tag. Use PDF remediation software to edit the Language tag.

The Language tag value must be an IETF BCP 47 language tag, which are codes applied in a vast range of computer systems to identify human languages. Each primary language, linguistic region and dialect is identified using a unique code.

Table 2. Example languages and language tags

Language and Linguistic Region/Dialect	IETF BCP 47 Language Tag
English (Australian)	en-AU
English (Great Britain)	en-GB
English (United States)	en-US

In most cases, a language tag will consist of a primary language tag and a linguistic region subtag in the following format: xx-YY

The Internet Assigned Numbers Authority (IANA) maintain a Language Subtag Registry on their website:

<https://www.iana.org/assignments/language-subtag-registry/language-subtag-registry>

You can perform a text based search on the registry to find language and region subtags that can be combined using a hyphen to form a complete language tag.

Example

German (language) subtag: de

Austria (region) subtag: AT

German (Austrian) language tag: de-AT

3.3.3. Role Mapping

Your document may contain non-standard structure tags in addition to the standard structure tags present in Tagged PDFs (refer to Appendix D for a complete list of standard structure tags). All non-standard structure tags eventually link to standard structure tags. The link between a non-

standard structure tag and another structure tag is known as role mapping.

This section briefly explains role mapping and how to resolve role mapping related errors.

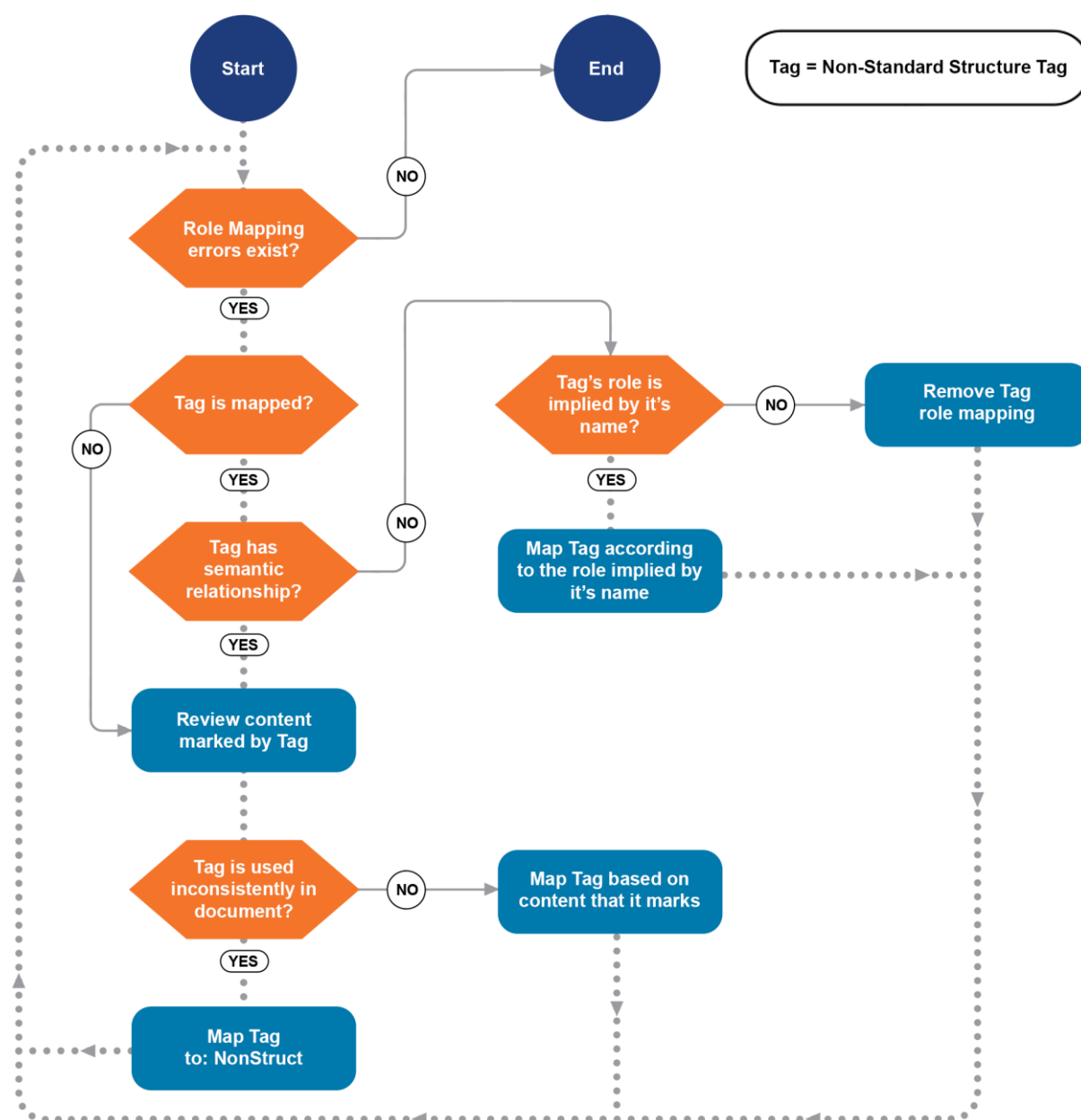


Figure 10. How to repair PDF role mapping errors

A non-standard structure tag is a structure tag that is not defined in PDF standards. Word to PDF conversion software may output non-standard structure tags to mark content in your document that share a common

Word style or purpose. Since non-standard structure tags are not defined in standards, they have no semantic value of their own. By mapping a non-standard structure tag to a standard structure tag, software is able to understand the role of the non-standard structure tag in your document.

PDF remediation software that provides the ability to edit role mappings will display a list of non-standard structure tags and the structure tags that each non-standard structure tag is mapped to.

Non-standard structure tags may be mapped to other non-standard structure tags, however in most cases they are mapped to standard structure tags.

Table 3. Example role mappings

Non-Standard Structure Tag	Structure Tag (Role)
Diagram	Figure
Footnote	Note
Heading1	H1

Ideally, non-standard structure tags should ultimately be mapped to standard structure tags that are appropriate for describing the content that the non-standard structure tags mark in your document.

If role mapping errors exist, it will be necessary to review and edit the role mappings in your document. Common role mapping errors are:

- Standard structure type "Y" is remapped to "X".
- Structure type "X" is mapped in a circular fashion.
- Non-standard structure type "X" is neither mapped to a standard structure type nor to another non-standard structure type.

(X and Y will be the names of the non-standard structure tag and standard structure tag)

If there are no role mapping errors in your document, skip this section, otherwise follow these instructions:

1. Open your PDF remediation software's role mapping editing form.
2. Find the non-standard structure tags that were mentioned in the role mapping error messages.
3. For each non-standard structure tag that caused errors, consider if the name of the non-standard structure tag is a hyponym of the standard structure tag it maps to. For instance, a Diagram is a type of Figure, and a Footnote is a type of Note.
 - a. If there is a semantic relationship between the non-standard structure tag and the structure tag it maps to; or the non-standard structure tag is not mapped:
 - Find instances in your document where the non-standard structure tag is marking content.
 - Map the non-standard structure tag to a standard structure tag that would normally contain the content marked by the non-standard structure tag. e.g. Map the non-standard structure tag to H1 if it being used to mark top-level headings.
 - If the non-standard structure tag is being used inconsistently throughout your document, map the non-standard structure tag to: NonStruct
Inconsistent use of non-standard structure tags are typically caused by technical limitations in Word to PDF conversion software.
 - b. If you find that there is no semantic relationship between a non-standard structure tag and the structure tag it maps to, follow these steps to correct the role mapping:
 - If the non-standard structure tag value implies that the tag is used for bodies of text, map it to the standard structure tag: P
 - If the non-standard structure tag value implies that the tag is used for words that are not part of a paragraph, map it to the standard structure tag: Span

- If the non-standard structure tag value implies that the tag is used for images, map it to the standard structure tag: Figure
- If the non-standard structure tag value implies that the tag is used for headings, map it to the standard structure tag: H1 to H6 (depending on the implied heading level)
- If the non-standard structure tag value implies that the tag is used for a section of a document such as the Header and Footer, map it to the standard structure tag: Sect
- In all other cases, remove the non-standard structure tag's role mapping and go to step 3a.

4. Close the role mapping editing form

3.3.4. Logical Structure

The logical structure of your document consists of structure tags arranged in a tree structure and sorted by reading order. PDF remediation software may represent your logical structure as a navigable tree. This section describes error messages you may encounter regarding your document's logical structure and provides methods to address each error.

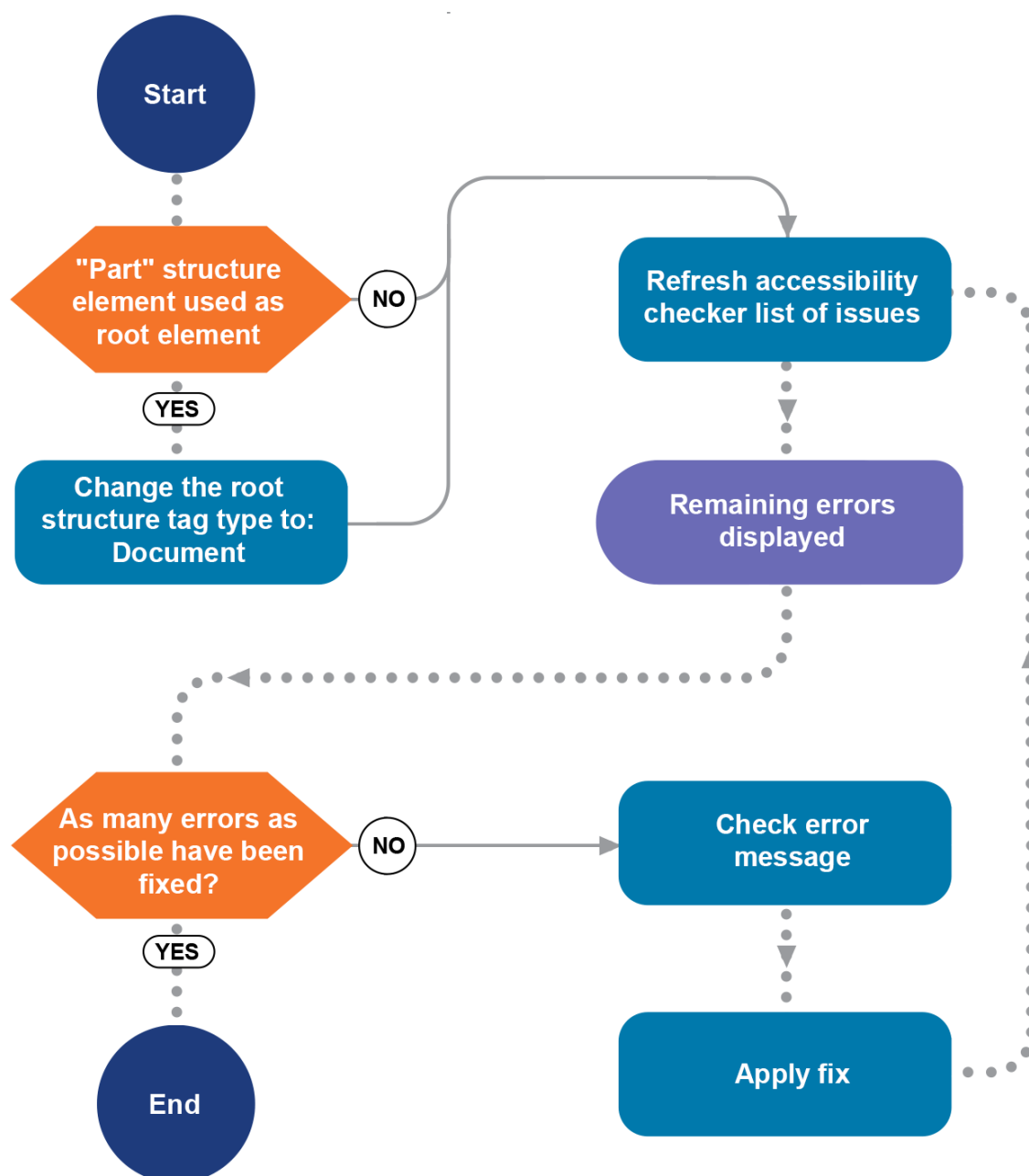


Figure 11. How to repair PDF logical structure errors

"Part" structure element used as root element

PDFs output by Microsoft Word 2013 and 2016 will usually contain this error. This error occurs when the structure type of the root structure tag in the logical structure tree is "Part". To resolve this issue, change the structure type of the root structure tag to: Document.

Alternative text missing for “Figure” structure element

If this error occurs, one or more Figure structure tags are missing alternative text. In order to resolve this issue, each Figure structure tag that raised an error should be checked.

- If the tagged figure is decorative, change the assigned Figure structure tag type to: NonStruct
- If the tagged figure is non-decorative, enter a concise description of the image into the Alternative Text property of the assigned Figure structure tag.

Alternative description missing for an annotation

This error will occur when Link structure tags are missing a value in their Link Annotation Contents property. In order to resolve this issue, each Link Annotation that raised an error should be checked.

If the Contents property of a Link Annotation is empty, enter a concise description of the linked annotation into the Contents property.

"Figure" element on single page with no bounding box

This error will occur when Figure structure tags are missing a Bounding Box attribute within their attribute object properties. This error can be resolved by adding Bounding Box attributes to affected Figure structure tags.

If a Figure structure tag exists for decorative purposes only, change the structure tag type to: NonStruct

Changing the structure type to NonStruct will eliminate the need to have a Bounding Box defined for the Figure.

In all other cases, proceed to editing the Figure structure tag's properties.

A Bounding Box is an invisible rectangular shape that defines the borders of a Figure using 4 values: Left, Bottom, Right, Top (in this order)

When entering values for a Bounding Box, be mindful of the value ranges described below.

Table 4. Bounding box value characteristics

Value	Description	Range on an A4 page	Range on an A3 page
Left	X axis value of the left side of the border	0 to 596	0 to 842
Right	X axis value of the right side of the border	0 to 596	0 to 842
Top	Y axis value of the top side of the border	0 to 841	0 to 1190
Bottom	Y axis value of the bottom side of the border	0 to 841	0 to 1190

The Right value should be greater than the Left value and the Top value should be greater than the Bottom value.

The X axis from Left to Right will start at 0 and end at the maximum value.

The Y axis from Top to Bottom will start at the maximum value and end at 0.

Note that the maximum value can be determined by referring to the Bounding Box Value Characteristics table.

Ensure that the Bounding Box values define a rectangle that tightly surrounds the Figure on the page.

PDF remediation software may provide a visual representation of the Bounding Box over the Figure which can be used to fine tune the values entered.

Possibly inappropriate use of a "X" structure element

This error will usually be displayed if structure tags have been assigned placement layout attribute values within their properties that are inconsistent with their default placement layout type. This error will also be raised when a structure tag is not being used for its purpose.

There are 4 possible values for placement layout attributes:

- Block
- Inline
- Before
- Start
- End

Inline and Block are the most frequently used placement layout attribute values.

Table 5. Common standard structure tags and their default placement layout attribute values

Standard Structure Tag	Placement
Figure	Inline
Span	Inline

Standard Structure Tag	Placement
P	Block
H1 to H6	Block
Table, TR, TH, TD	Block
Note	Block
Div	Block

If a Figure structure tag's placement layout attribute value is set to Block, it would cause an error to be generated as the value does not match the expected value (Inline).

To resolve this error, change the placement layout attribute value for the affected structure tag to Inline if the current value is Block or to Block if the current value is Inline. If no value had been set, change the placement layout attribute value to Block.

If the error persists, check that the structure tag type is being applied correctly for the content it marks. For instance, if the structure tag contains an image and its tag type is not Figure, change the structure tag type to: Figure

If the structure tag is being used for decorative purposes, change the structure tag type to: NonStruct

Invalid use of a "Span" structure element

This error is raised when a Span structure tag is not used to tag text content or exists on a level in the logical structure tree that cannot

contain Span structure tags according to standards. Bugs/technical limitations in Microsoft Word are usually responsible for this error.

Locate each Span structure tag causing this error and change their structure tag type to: NonStruct

Invalid use of a "TBody" structure element

This error is raised when a TBody structure tag is not accompanied with a sibling THead structure tag. Bugs/technical limitations in Microsoft Word are usually responsible for this error.

A table with a top header row and no other headers should consist of the following structure tags:

- Table
 - THead
 - TR
 - TH
 - TBody
 - TR
 - TD
 - TFoot (optional)
 - TR
 - TD

If a THead structure tag is missing from the Table structure tag, the TBody structure tag becomes redundant.

Locate each TBody structure tag causing this error and change their structure tag type to: NonStruct

Table header cell has no associated subcells

This error will occur when TH structure tags are missing a scope table attribute value. A TH structure tag represents a header cell in a table.

Check each TH structure tag that is causing errors and if their scope table attribute value is not set, enter a value as follows:

1. If the TH structure tag is the header for a table column, change the scope table attribute value to: Column
2. If the TH structure tag is the header for a table row, change the scope table attribute value to: Row
3. If the TH structure tag is the header for a table column and row, change the scope table attribute value to: Both

All other Logical Structure errors

For any other logical structure error, consider whether the structure tags raising errors add semantic value to the document. If they do not, change their structure tag type to: NonStruct

You may attempt to change the structure tag type of a tag that is causing errors if you feel that there is a more appropriate tag type to mark the content that it is related to. To help gain an insight into the purpose of each standard structure tag and whether or not they may be appropriate for the content being marked, refer to Appendix D.

3.3.5. Finalising

If the only remaining issue present in your document is *PDF/UA identifier missing*, it means that your document does not contain an indicator to identify it is as PDF/UA compliant.

The PDF remediation tool that you are using may provide a shortcut to add the PDF/UA identifier for you automatically, otherwise you will need to add it manually by editing your document's XMP metadata as the identifier is in XMP format.

Please do not add the identifier to your document unless all other accessibility issues have been resolved. Before you begin adding the identifier manually, note that these steps involve searching for text,

inserting text and replacing text within XMP metadata which is a text-based format:

Look for text in the document's XMP metadata that starts with

`<rdf:Description` and includes the text

`xmlns:dc="http://purl.org/dc/elements/1.1/"`

Examples of how matching text may appear are available in Appendix A.

If you were able to locate a block of text matching the aforementioned criteria:

1. Append the text

`xmlns:pdfuaid="http://www.aiim.org/pdfua/ns/id/"` on a new line immediately before `xmlns:dc="http://purl.org/dc/elements/1.1/"`

2. Append the text `<pdfuaid:part>1</pdfuaid:part>` before `</rdf:Description>`

If you were unable to locate any of the above text, your document will not be able to be PDF/UA compliant as it is missing XMP metadata.

Save changes to your document (if any).

If you have repaired all issues, then you will have achieved generating a PDF/UA compliant PDF. Otherwise, your PDF will not be compliant but it may be more accessible than it was prior to being processed using this section of the guide.

Appendix

Appendix A

Example 1

```
<rdf:Description rdf:about=""
  xmlns:pdf="http://ns.adobe.com/pdf/1.3/"
  xmlns:xmp="http://ns.adobe.com/xap/1.0/"
  xmlns:xmpMM="http://ns.adobe.com/xap/1.0/mm/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  pdf:Producer="Acrobat Distiller 4.0 for Windows"
  pdf:Keywords="Test, experimental, sample"
  xmp:ModifyDate="2019-12-12T09:43:11+09:00"
  xmp:CreateDate="2001-02-07T12:43:33Z"
  xmp:MetadataDate="2019-12-12T09:43:11+09:00"
  xmpMM:DocumentID="uuid:e078b573-fbb6-4975-ac4e-7ad8f5315af3"
  xmpMM:InstanceID="uuid:85cb8a4b-48ea-4f72-92e8-78b01658b87f"
  dc:format="application/pdf">
```

Example 2

```
<rdf:Description rdf:about=""
  xmlns:dc="http://purl.org/dc/elements/1.1/">
```

Appendix B

```
<dc:title>
  <rdf:Alt>
    <rdf:li xml:lang="x-default">ENTER DOCUMENT TITLE
    HERE</rdf:li>
  </rdf:Alt>
</dc:title>
```


Appendix C

```
<?xpacket begin="" id="W5M0MpCehiHzreSzNTczkc9d"?>
<x:xmpmeta xmlns:x="adobe:ns:meta/" x:xmptk="3.1-701">
  <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-
ns#">
    <rdf:Description rdf:about=""
xmlns:dc="http://purl.org/dc/elements/1.1/">
      <dc:title>
        <rdf:Alt>
          <rdf:li xml:lang="x-default">ENTER DOCUMENT TITLE
HERE</rdf:li>
        </rdf:Alt>
      </dc:title>
    </rdf:Description>
  </rdf:RDF>
</x:xmpmeta>
<?xpacket end="w"?>
```

Appendix D

The following tables briefly describe all structure tags/elements that are defined in ISO 32000-1 (PDF 1.7).

Container Tags

Container tags mark major content groups. The first tag in all documents should be the Document tag as it is expected to contain all other tags.

Standard Structure Tag	Name	Role/Purpose
Document	Document	Represents an entire document. Only one Document structure tag should appear in a document.
Part	Part [of Document]	Partitions content in a document.
Art	Article	Defines an article, such as a blog post within a document.
Sect	Section	Identifies sections within a document, part of a document or an article inside a document.
Div	Division	Groups content without impacting semantics.

General Tags

General tags are used to mark text, image and interactive content.

Standard Structure Tag	Name	Role/Purpose
BlockQuote	Block Quotation	Encloses a quotation inside a standalone block.
Code	Code	Defines a snippet of computer language.
Figure	Figure	Defines non-textual visual content such as charts and photos.
Form	Form	Defines an interactive data entry form.
Formula	Formula	Marks an equation or formula.
Link	Link	Defines an interactive link to content within a document or to online content.
P	Paragraph	Contains a paragraph of text.
Quote	Quotation	Encloses an inline quotation.
Span	Span [of Text]	Contains inline text.

Heading Tags

Heading tags mark heading text. Headings are a key instrument for defining a document's structure.

Standard Structure Tag	Name	Role/Purpose
H	Heading	Represents a heading in a document or subdivision. It should not be used in documents with multiple heading levels.
H1	Heading Level 1	Denotes a 1st level heading.
H2	Heading Level 2	Denotes a 2nd level heading.
H3	Heading Level 3	Denotes a 3rd level heading.
H4	Heading Level 4	Denotes a 4th level heading.
H5	Heading Level 5	Denotes a 5th level heading.
H6	Heading Level 6	Denotes a 6th level heading.

List Tags

List tags mark lists such as bullet point or numeric lists.

Standard Structure Tag	Name	Role/Purpose
L	List	Defines a list.
LI	List Item	Defines an item in a list.
LBody	List Body	Encloses content that pertains to a list item.

Table Tags

Table tags work collectively to mark tables, including their header, body and footer sections. Tables without headers or footers should not have a table body section marked since the rows are implied to be the table body.

Standard Structure Tag	Name	Role/Purpose
Table	Table	Defines a table, consisting of columns and rows.
THead	Table Header	Denotes the section of a table that contains the table's header content.

Standard Structure Tag	Name	Role/Purpose
TBody	Table Body	Denotes the section of a table that contains the table's main content.
TFoot	Table Footer	Denotes the section of a table that contains the table's footer content.
TR	Table Row	Defines a table row inside a table.
TH	Table Header Cell	Defines a single table cell that contains either a column heading or a row heading.
TD	Table Cell	Defines a single table cell that contains table content.

Referencing and Indexing Tags

Referencing and indexing tags mark annotations, labels, notes and lists containing internal indexes.

Standard Structure Tag	Name	Role/Purpose
Annot	Annotation	Marks annotations that are not links or widgets, such as audio or video content.

Standard Structure Tag	Name	Role/Purpose
BibEntry	Bibliography Entry	Identifies individual entries in a bibliography.
Caption	Caption	Contains text describing a list, table or figure.
Index	Index	Groups document indices, usually consisting of a list and reference items.
Lbl	Label	Identifies content that labels other content. e.g. It may contain a number that numbers an item in a numeric list.
Note	Note	Encompasses explanatory text such as footnotes and endnotes.
Reference	Reference	Contains content that refers to other content within a document.
TOC	Table of Contents	Defines a table of contents.
TOCI	Table of Contents Item	Defines an item in a table of contents.

Special Tags

Special tags are used to hide content in a document's logical structure. Content marked by special tags will be ignored by screen readers.

Standard Structure Tag	Name	Role/Purpose
NonStruct	Non-Structural [Element]	Makes a tag invisible in a document's logical structure. If it contains child tags, the child tags will remain visible.
Private	Private	Makes a tag and all its child tags invisible in a document's logical structure.

Ruby Annotation Tags

You may encounter ruby annotations - also known as rubi and ruby characters - in documents containing Asian text, in particular Japanese and Korean text. Ruby annotations are concise annotations displayed in small font that appear directly above text being annotated.

The following image demonstrates ruby annotation layout on the right, and two examples using German characters on the left. The top line in the image contains the ruby annotation text and the bottom line contains the ruby base text.

Pronunciation	ae	ss
Character	ä	ß

It is safe to disregard ruby annotations as their absence will not prevent documents from passing accessibility checks. If you are familiar with

Asian characters and pronunciations, tag Asian characters and their respective pronunciations where appropriate using the ruby standard structure tags as this may benefit your target audience.

Standard Structure Tag	Name	Role/Purpose
Ruby	Ruby Annotation	Defines a ruby/rubi annotation such as Furigana and Bopomofo (Zhuyin).
RB	Ruby Base Text	Encloses text that is explained using Ruby Annotation Text.
RT	Ruby Annotation Text	Encloses text that explains content in a Ruby Base Text.
RP	Ruby Parenthesis	Encloses parenthesis inside Ruby Annotation Text. Parenthesis appear in cases where software does not fully support Ruby.

Warichū Annotation Tags

You may encounter warichū annotations in documents containing Japanese or other East Asian text. Warichū annotations are annotations that are split into two lines that occupy the space of a single line of text, typically wrapped in parenthesis or spaces.

The following images demonstrate an example of warichū annotation using the pangram "The quick brown fox jumps over the lazy dog". The

words “very fast and agile” have been added as a warichū annotation for the word “quick”.

The quick (very fast and agile) brown fox jumps over the lazy dog

d
o
g

f
o
x
j
u
m
p
s
o
v
e
r
t
h
e
l
a
z
y

T
h
e
q
u
i
c
k
(
v
e
r
y
f
a
s
t
)
b
r
o
w
n

It is safe to disregard warichū annotations as their absence will not prevent documents from passing accessibility checks.

Standard Structure Tag	Name	Role/Purpose
Warichu	Warichū	Defines a Warichū annotation.
WT	Warichū Text	Encloses the text of a Warichū annotation.
WP	Warichū Parenthesis	Encloses parenthesis or spaces that surround Warichū Text.