



Meeting the Standard: How do I know if my PDF file is compliant?

Chad Chelius

www.accessibilityunraveled.com

chad@chaxtc.com

linkedin.com/in/chadchelius





PDF/UA Compliance

The PDF/UA standard was designed to help users and software developers create documents and applications to a known standard to provide a common goal for users to follow to ensure that PDF files that we create are done to the highest level of accessibility. With a known and reliable standard we can ensure that our documents will be compliant and provide a good user experience to users of assistive technology.

Accessibility is a series of automated and manual checks that need to be performed to ensure that our PDF documents are accessible. The automated side is easy, the manual side on the other hand relies on knowledge and skill of the user remediating the file to ensure that the document is fully accessible.

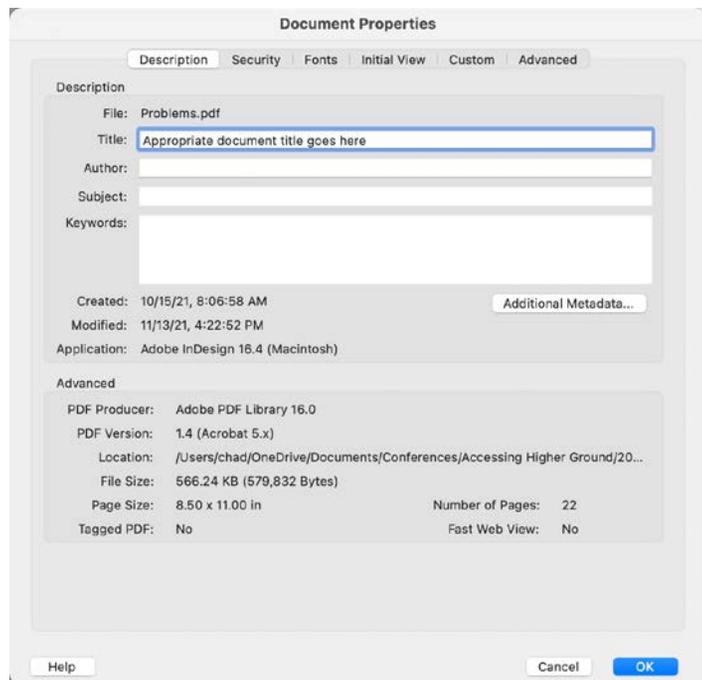
The automated portion of the equation can be performed using a number of different applications. But only a few will check a document against the PDF/UA standard. The PAC 2024 checker which was released earlier in 2024, is arguably the defacto standard accessibility checker in use to verify document accessibility to the PDF/UA and WCAG standards. The PAC 2024 checker does a fantastic job of checking for problems in a file and an even better job of showing you where exactly the problem is occurring in the document.

The errors themselves however, seem to require a translator to understand exactly what the error is. As good as it is at showing you what element in the document is causing the error, understanding how to fix it can be tricky to say the least. In this session I hope to explain what some of those errors mean and show you how to address the problem in Adobe Acrobat Pro. Although there are 3rd party applications that can make the process of addressing these errors a bit easier, Adobe Acrobat Pro is a powerful tool that can be used to address most of the problems that you're likely to encounter when checking your document using the PAC 2024 checker.

Error: Title in XMP Metadata

Every accessible PDF file needs to contain a document title. When a title doesn't exist in the file, you'll receive this error.

Fix: Most source applications that provide the ability to output an accessible PDF file, provide the ability to add a title that will be included in the exported PDF file. So it's preferable to add the title there. However in Adobe Acrobat, you can easily add a title by choosing File > Properties and clicking on the Description tab and entering a title into the Title field.

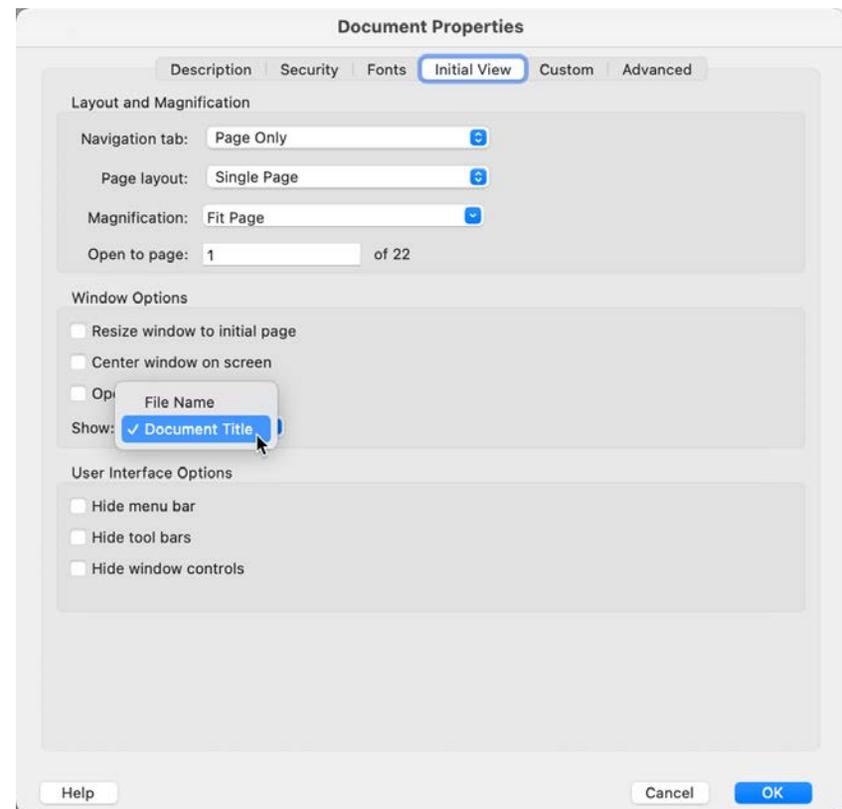


Adding an appropriate document title to the document.

Error: Display of Document title in window title

Although it's true that every accessible PDF file needs a document title, the secondary requirement is that the title needs to be displayed when the PDF file is being viewed.

Fix: This error can be fixed by choosing File > Properties, clicking on the Initial View tab and setting the Show drop-down menu to Document Title.

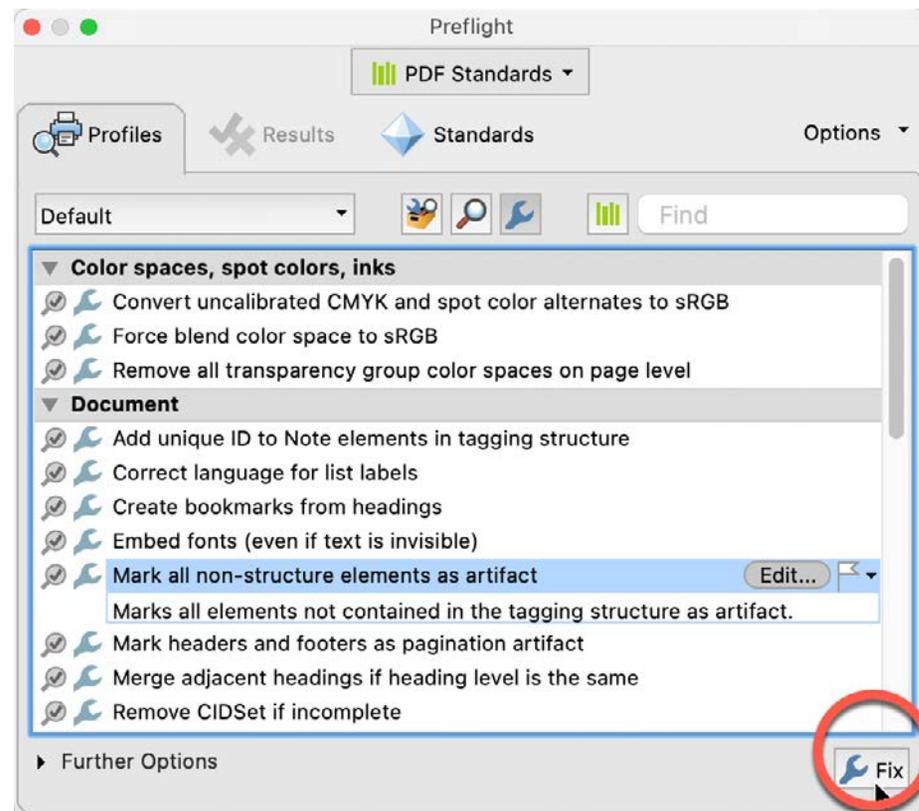


Telling the document title to be displayed when the document is open.

Error: Path Object not Tagged

Every element in a PDF file needs a tag applied to it. It doesn't matter what it is, even a path needs to have a tag applied. In the case of a path it will likely need to be tagged as an artifact but nevertheless it needs to be tagged. Many source applications neglect to tag extraneous paths appropriately. These paths are often in the form of an underline under a hyperlink, or a shape drawn on the page. One common culprit is when a table has borders around the cells. Every side of every cell is often an individual path and those paths are often untagged. It can be scary to run a file through the PAC 2021 checker to see 1,576 errors in the document. This error is actually scarier than it looks as those tiny little paths can really add up.

Fix: Acrobat has an automated way to fix these errors and it lives in the Preflight panel. Open the Preflight panel by clicking on the Print Production tool in the Tools list or you can just search for Preflight in the search tools area at the top of the tools list. With the Preflight panel open, make sure that PDF Standards is chosen from the drop-down menu at the top of the window, then click on the wrench icon  just below that drop-down menu. In the Document category of the main window, click on the "Mark all non-structure elements as artifact"  fix and click the Fix button  in the lower-right corner of the window. You'll need to save your file as a new file or just overwrite the existing one and when the Fix-up is finished, it will let you know how many items it found and fixed.

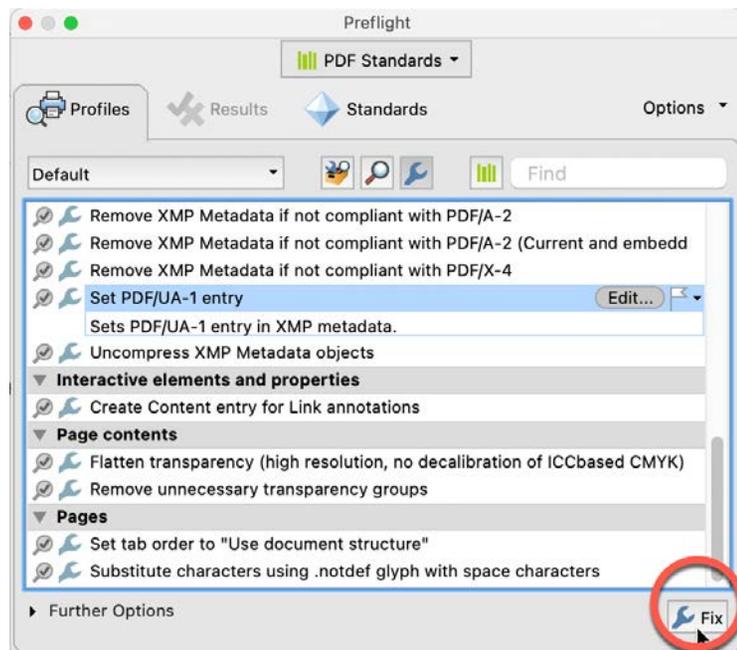


Running the "Mark all non-structure elements as artifact" fix-up.

Error: PDF/UA Identifier Missing

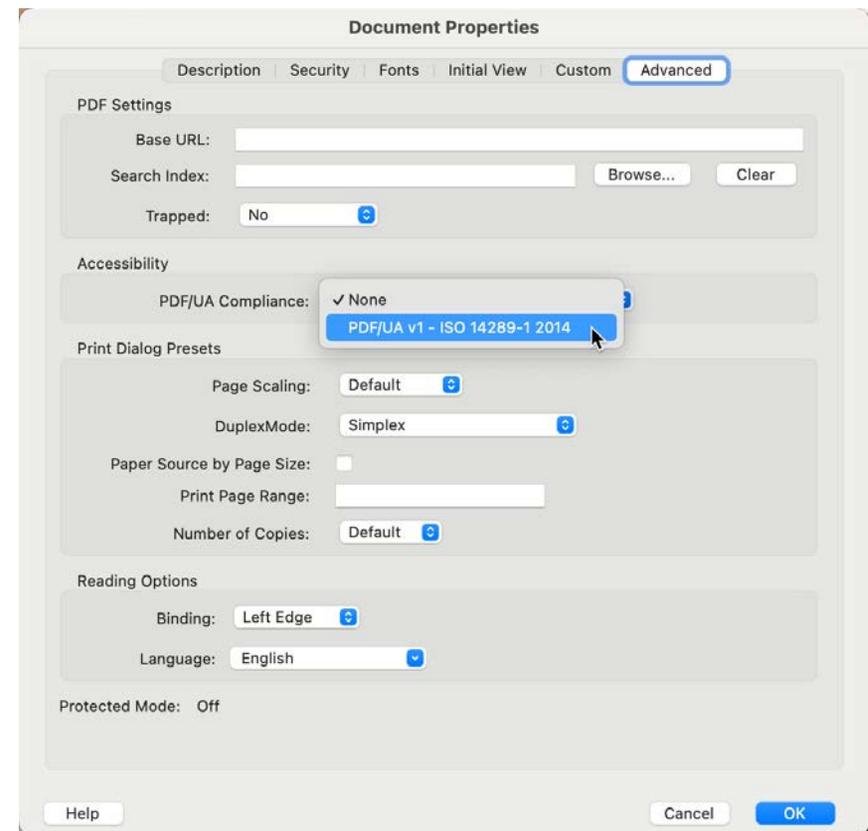
For a PDF/UA compliant file, an identifier needs to be added to the file that identifies the file as being PDF/UA compliant. This is a manual step that needs to be performed by the person remediating the file. I like to perform this step as a final step in the remediation process once I'm sure the file meets all of the PDF/UA requirements.

Fix: In Adobe Acrobat, open the Preflight panel and in the Document Info and Metadata category, choose the fix-up called Set PDF/UA-1 entry. This will add the entry to the file indicating that the file is PDF/UA compliant.



Setting the PDF/UA-1 identifier.

Additionally, Adobe Acrobat has made it even easier to add the PDF U/A identifier to a document by opening up the Properties panel in Adobe Acrobat. With the Properties panel open, click on the Advanced tab. In the Accessibility section of the window, you'll find a PDF/UA Compliance drop-down menu. Simply choose PDF/UA-1 from the drop-down menu and the identifier will be added to the file.

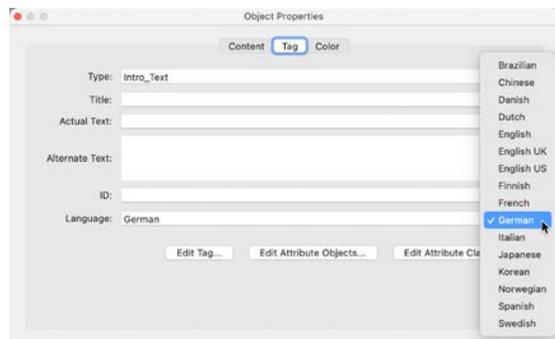


Adding the PDF/UA identifier in the Document Properties of Adobe Acrobat.

Error: Natural language of objects

For a PDF to be compliant, it's important that the appropriate language of text elements is defined so that users of assistive technology will know which language the text is and can be read appropriately provided the user has the localized version of that language installed on their computer. Just because text is defined as a certain language, doesn't mean that a user's assistive software will read it in that language. There's more to it than that. They need the localized version of that language installed to provide that capability.

Fix: Set the language appropriately. The language can be set in a PDF file at the document level by choosing File > Properties in Adobe Acrobat, clicking on the Advanced tag, and choosing the appropriate language from the Language drop-down menu or by typing in the appropriate language code. Furthermore, the language can be defined at the tag level by selecting a tag in the Tags pane, right-clicking and choosing Properties and choosing the appropriate language from the Language drop-down menu.

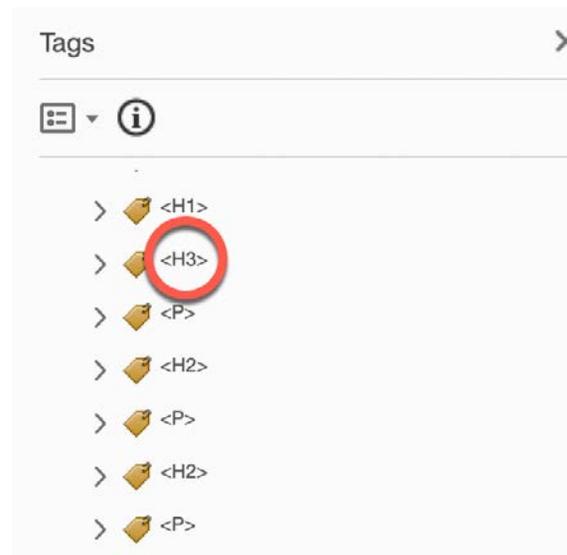


Choosing the appropriate language from the Language drop-down menu.

Error: Nesting of Heading Levels

Heading levels in a PDF file provide users of assistive technology with the ability to navigate the sections of the document. When using heading levels, it's important that the sequence of headings is logical (i.e. H1, H2, H3). If the headings in a document don't begin with an H1 or skip sequence (i.e. H1, H2, H4), then you'll receive the Nesting of Heading Levels error.

Fix: There's no automated fix for this. Essentially you need to review the document by walking the tags tree to identify when heading levels have been skipped. The fix is to rename the tags either manually or by opening the properties panel for the tag and changing the tag name there.

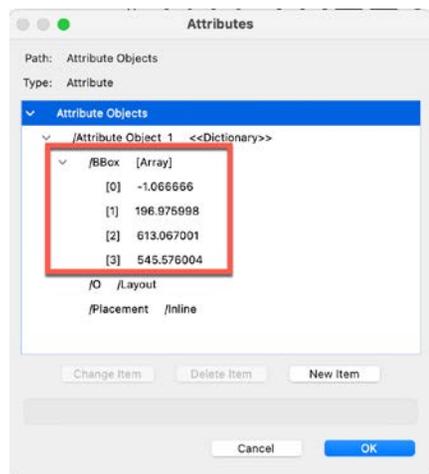


If heading tags do not occur in sequence, you'll receive the Nesting of Heading Levels error.

Error: Figure element on a single page with no bounding box

The blame for this one usually lies within the source application that created the file. Essentially, every figure in a PDF document needs to contain a BBox (bounding box) attribute that identifies the coordinates of the image on the page. When this attribute doesn't exist, you'll receive this error.

Fix: The fix for this in Acrobat isn't ideal, however it is possible. Start by artifacting the image in the Tags pane. Note that you may need to pull any Captions out of the Figure tag first and you may also want to copy any alternate text applied to the Figure prior to Artifacting it. Next, open the Reading Order panel, marquee the image, and click on the Figure tag in the Reading Order panel to re-tag the image as a Figure. This will properly define the BBox attribute for the Figure.



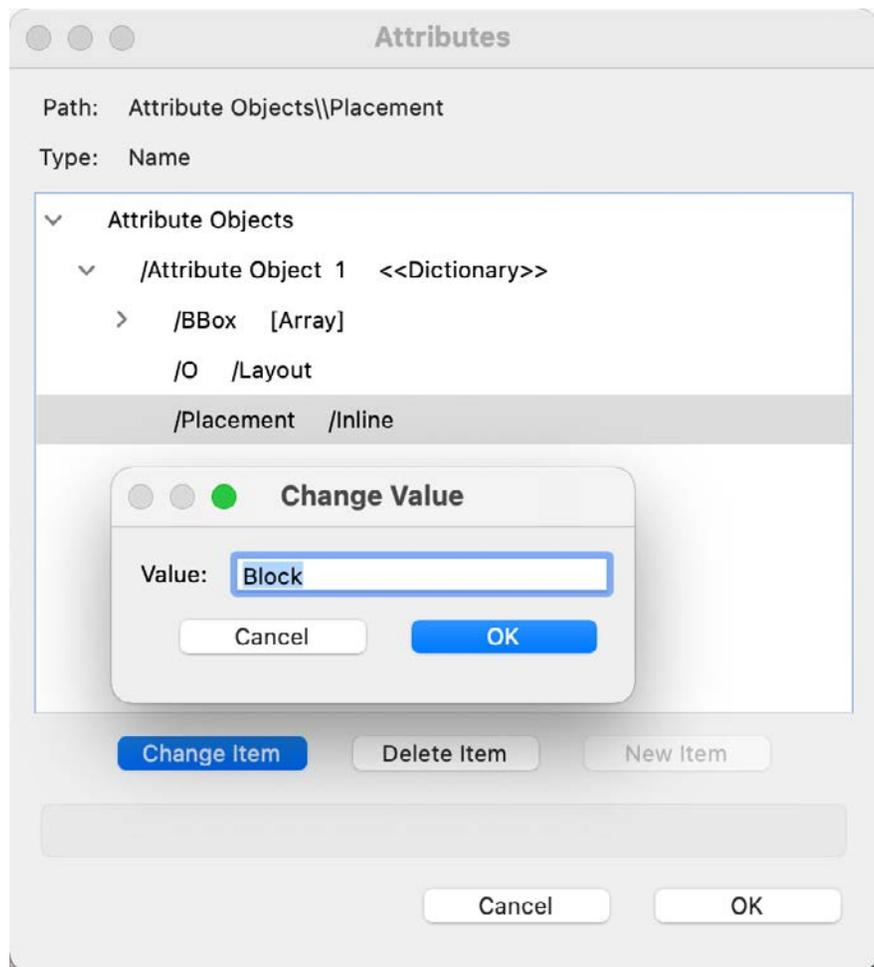
The BBox attribute is needed for all figures.

Error: Possible inappropriate use of an X structure element

The cause of this error is usually one of two things. Probably the most common cause of the error is that an element is in an illegal position within the Tags tree. Tags fall under two categories... Block-level and Inline-level. Inline tags within a block-level tags are natural and expected. But when a block-level element is nested within another block-level element, it will cause this error. The fix is typically to simply move the element out of its container. Common examples of this are a <Table> as a child of a <P> tag or a <P> tag as a child of a <Caption> tag.

The other cause of this error is similar to the first but slightly more challenging to fix. Sometimes tags can be defined as Block-level or Inline-level. In these cases, the placement attribute simply needs to be changed to accommodate its position in the Tag structure. For example a <Figure> at the root level of the Tags tree, needs to have its Placement attribute defined as Block instead of Inline. Conversely when a <Figure> is inline within a <P> tag, its placement attribute needs to be set to Inline.

Fix: To accomplish this, select the tag in question, then right-click on the tag and choose Properties. Click the Edit Attribute Objects button and open the resulting containers. If you see a Placement property, change it to Block or Inline. If no Placement property exists, create one by clicking on the New Item button and entering Placement for the Key and Block or Inline for the Value. Leave the value type set to Name.

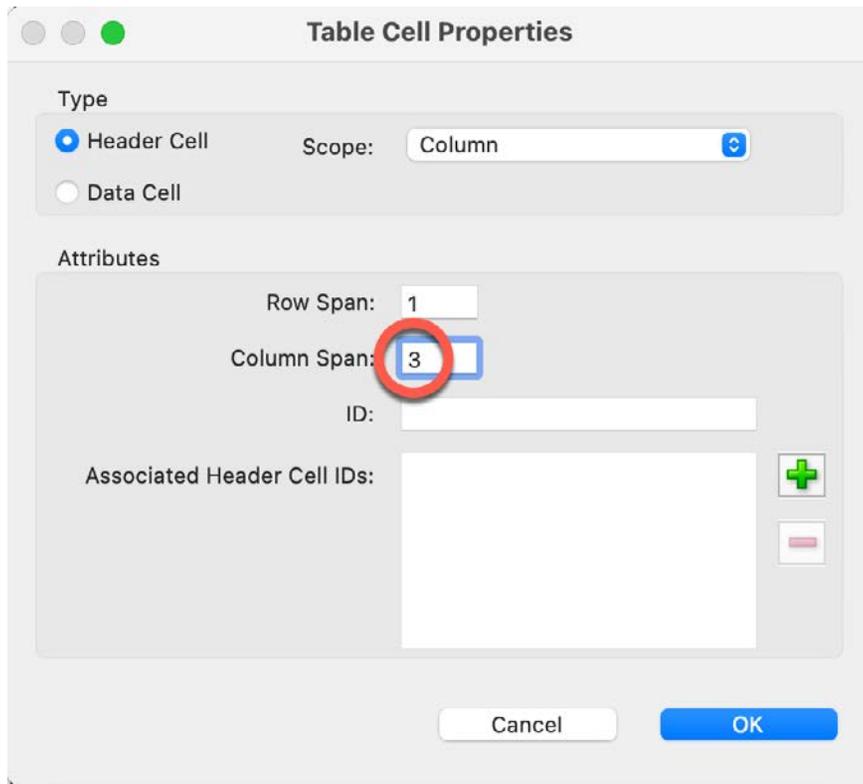


Changing the Placement attribute for a tag.

Error: Table regularity: Irregular table row

Tables rows in an accessible PDF file are all expected to have the same number of cells in each row. When that rule is broken, you'll receive a regularity error in both the Acrobat Accessibility checker and of course in the PAC 2021 checker. The fundamental cause of this error is when cells are merged in the source application. Some applications do a better job than others in properly defining these merged cells in the exported PDF file. It's incorrectly assumed by many that the fix is to simply avoid merging cells. This assumption is incorrect as this error can be resolved and the table properly defined in the PDF file.

Fix: To address this problem, a column span or row span needs to be defined for the offending cells in the PDF file. The easiest way to define this is by using the Table Editor in Adobe Acrobat. Open the Reading Order panel, click within the table you want to adjust, and then click on the Table Editor button in the Reading Order panel. You can now select the offending cell and right click on the cell and choose Table Cell Properties. In the Table Cell Properties dialog box, change the Column Span or Row Span property to the appropriate value. That value can be determined by looking at the row in the table that has the MOST number of cells. Use that as a reference and apply the appropriate span value to occupy the appropriate number of cells in the row or column that has the most number of cells.

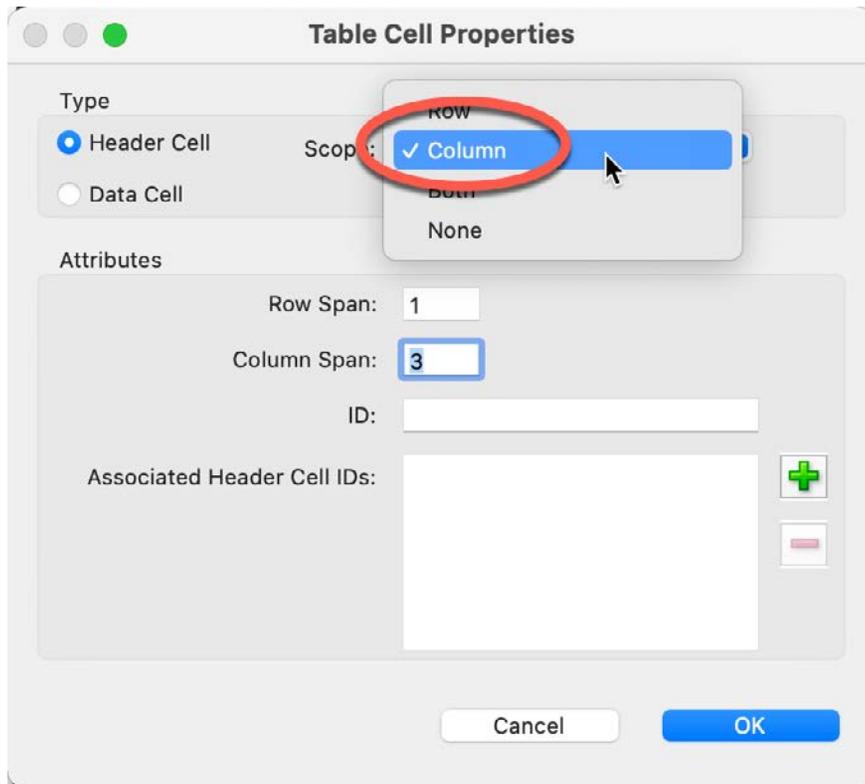


Defining a Column Span for a table cell.

Error: Table header cell has no associated subcells

This error occurs when your table contains header cells, but the scope has not been defined for those header cells. You can set the scope of a header cell to column, row, or both. Alternatively, you could also create header cell IDs and associate the data cells to the appropriate header cells. Defining scope for header cells is sufficient for simple tables with only one column or row header and is a fairly simple task in Adobe Acrobat where header cell IDs is required for more complex tables and is much harder in Adobe Acrobat.

Fix: This error can be addressed in the Table Editor within Adobe Acrobat. Open the Reading Order panel, click within the table you want to adjust, and then click on the Table Editor button in the Reading Order panel. You can now select a cell or range of cells and right click on any of the selected cells and choose Table Cell Properties. In the Table Cell Properties dialog box, choose a property (column, row, both) from the Scope drop-down menu.



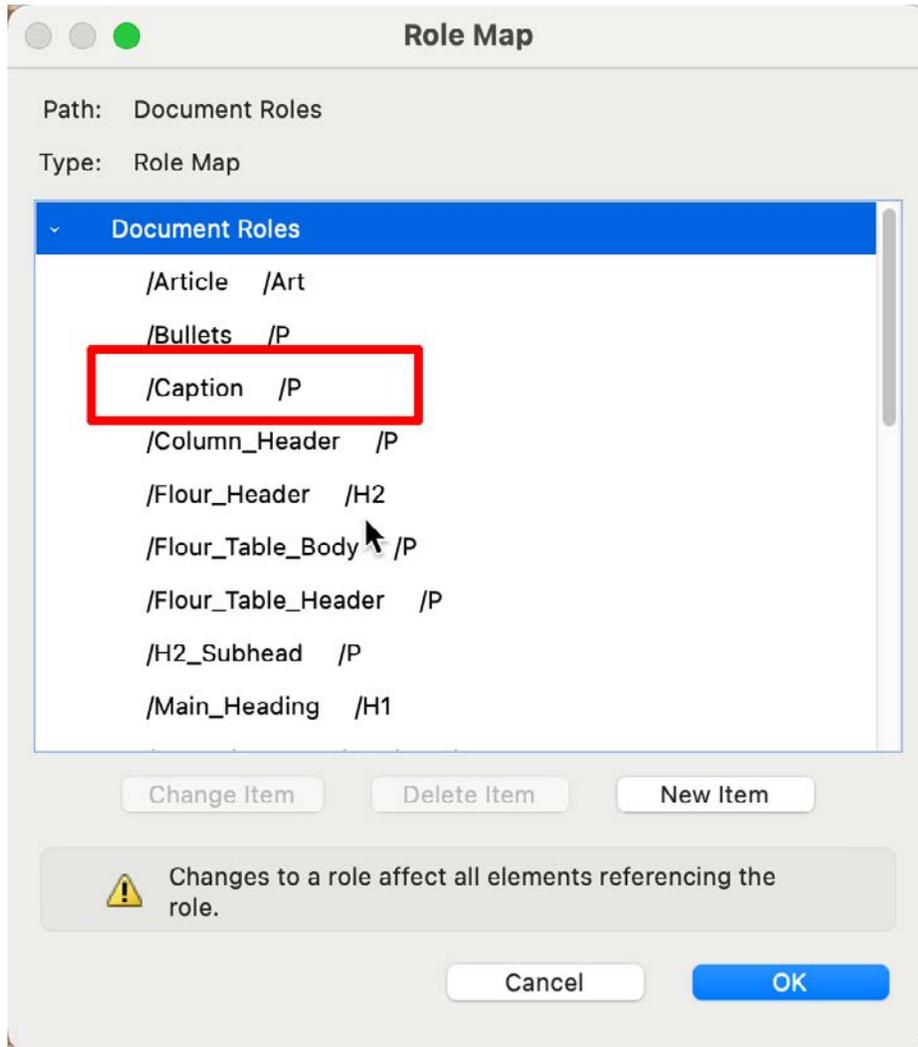
Choosing the Scope in the Table Cell Properties dialog box.

Error: Standard structure type X is remapped to Y

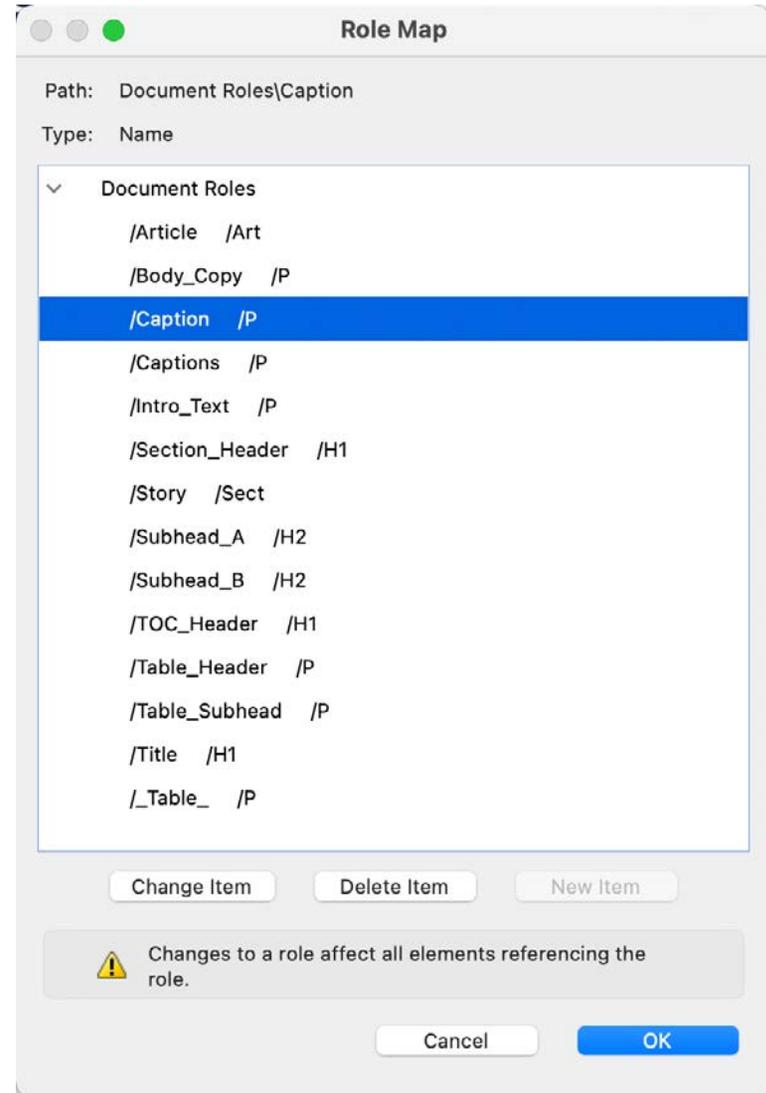
Role maps are created by the source application that creates the tagged PDF file. A role map is a feature that will map the name of the style in your source document, for example Body_copy to a standard PDF tag such as <P>. It's how the source application creates the connection to the PDF tag from the style used to format the text in the source document. The error above is caused when the name of the style in the source document is the name of an actual tag in the PDF file. For example creating a style called Caption and role mapping it to a <P> tag. Because both Caption and P are names of possible tags in the PDF file, this error occurs because you can't map a standard tag to another standard tag.

Fix: You have a couple of choices to address this problem but they both involve opening the same dialog box. Open the Tags pane and click on the Options button  at the top of the tags pane. From the list, choose Edit Role Map. In the Role Map dialog box, open the Document Roles category to view all of the role maps defined in the document. Find the role map that was causing the error and do one of two things:

1. Click the Change Item button to map the style to a more appropriate tag that is not a standard tag.
2. Delete the role map entirely if the style name is the name of a tag. The tag in the tags pane will now simply display as the name of that tag.



The role map dialog box in Adobe Acrobat showing a standard tag mapped to another standard tag.



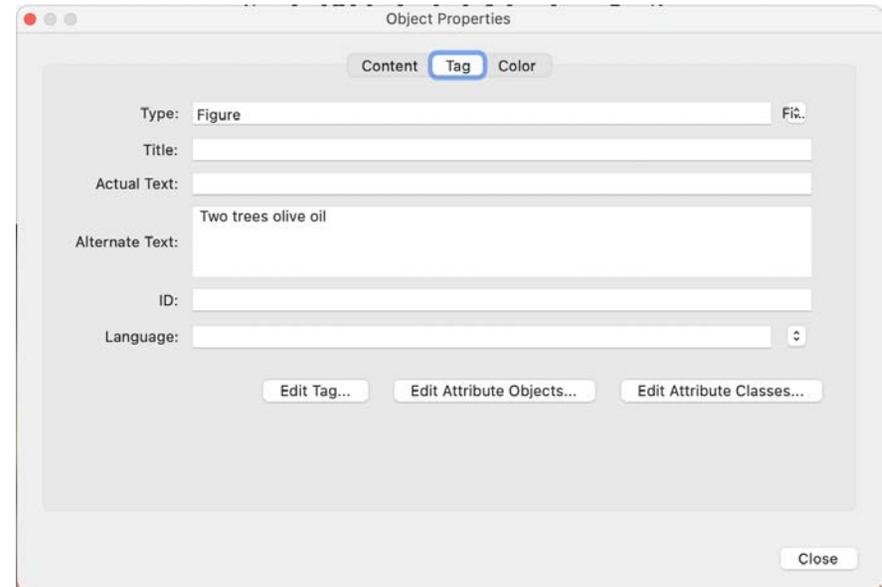
Adjusting the role map to avoid the PAC 2021 error.

Error: Alternative text missing for “Figure” structure element

We know that all images in an accessible PDF document that have value, need to have alternative text added to those images to describe them in their context. Without this alternative text added to images and graphics you’ll receive the above error when checking them using the PAC checker.

Fix: The fix for this is to simply add appropriate alternative text for the figures in your document. To add alternative text to a Figure in Adobe Acrobat, click on the <Figure> tag in the Tags pane, then right click on the tag and choose Properties. In the Object Properties dialog box, click on the Tag button at the top of the dialog box, then enter the appropriate alternative text in the Alternate Text field.

Alternatively, if the figure has no inherent value, you can artifact the figure by opening the <Figure> tag, selecting the content object, and right-clicking on the content object and choosing Change Tag to Artifact. Once you’ve artifacted the content object, you can delete the empty Figure tag that remains.



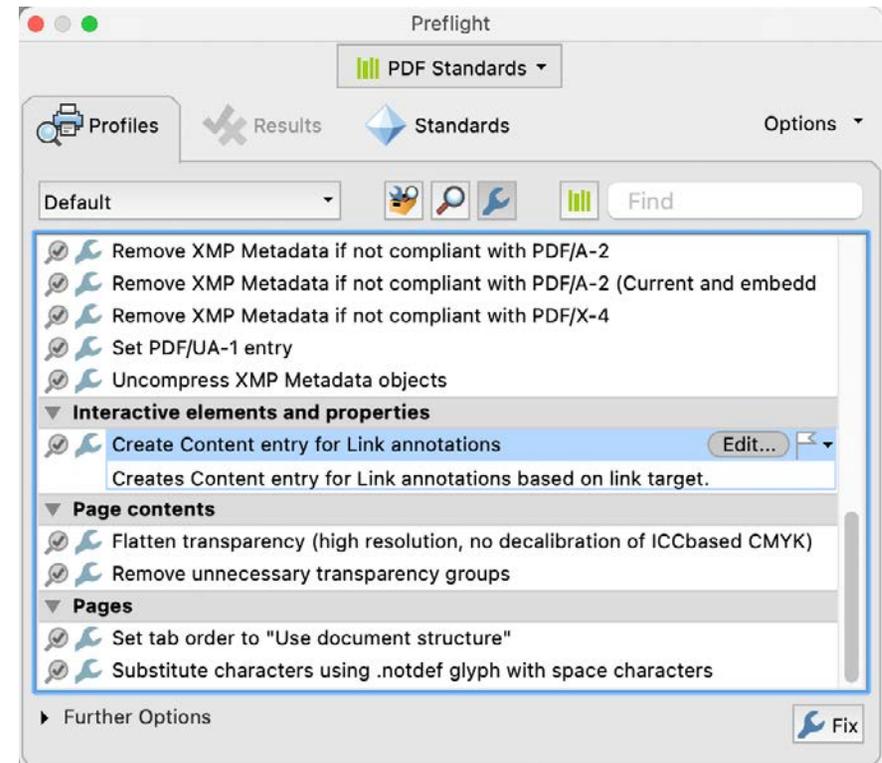
Applying Alternative text to a Figure in Adobe Acrobat.

Error: Alternative description missing for an annotation

This error often catches users by surprise. The idea that alternative text needs to be added to every link in a document can seem like a lot of work. The reality however is that this requirement stems from a period in time when the use of the words “click here” were used for hyperlinks in a document. This practice of course has been discouraged in favor of the use of more meaningful text such as “read the employee manual here” or “download the two trees employee manual”. When the text of the hyperlink is meaningful, fixing the above error becomes quite simple.

Fix: Adobe Acrobat allows us to address this error by applying a fix-up to automatically add the alternate text to every hyperlink of the document. How does it do this? Well provided you’ve used meaningful text for the hyperlinks in a document, the fix-up will take the text of the hyperlink and add that as the alternate text for the link. This is why it’s important that the links in your document are using meaningful text.

Open the Preflight panel in Adobe Acrobat. In the Interactive elements and properties category, run the Creates Content entry for Link annotations fix-up to set the alternative text for every link to the text that is linked in each instance.

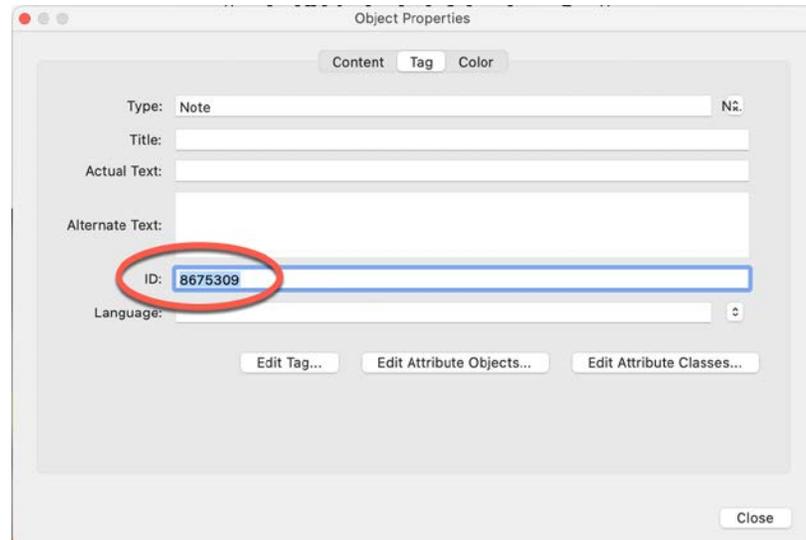


Running the Create Content entry for Link annotations fix-up in the Preflight panel.

Error: ID missing in “Note” structure element

When footnotes or endnotes are used in a document, both are tagged as <Note> structure elements. These notes are required to have a unique ID in each element to make each note unique in the document.

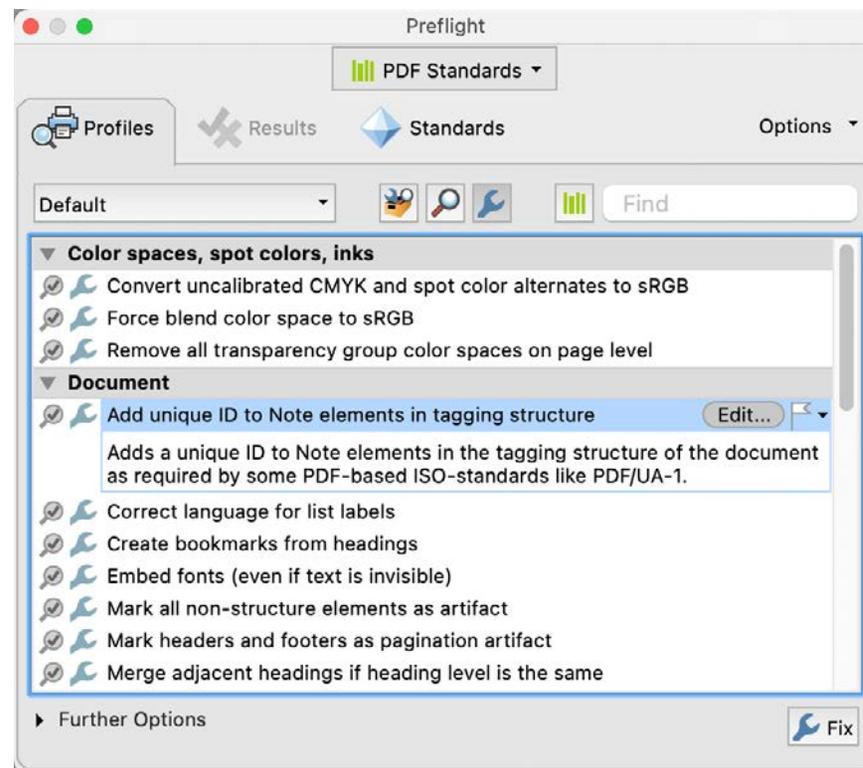
Fix: This task can be performed manually by opening up the Properties for the Note and entering a unique random number in the ID field.



Manually adding an ID to a Note element.

If you have a large number of <Note> elements in a document, adding the unique IDs manually is not desirable. Fortunately, Adobe Acrobat contains a preflight fix-up that can add unique IDs to <Note> elements automatically.

In Adobe Acrobat, open the Preflight panel and in the Document category, run the Add unique ID to Note elements in tagging structure fix-up to automatically add unique and random IDs to all <Note> structure elements in the document.

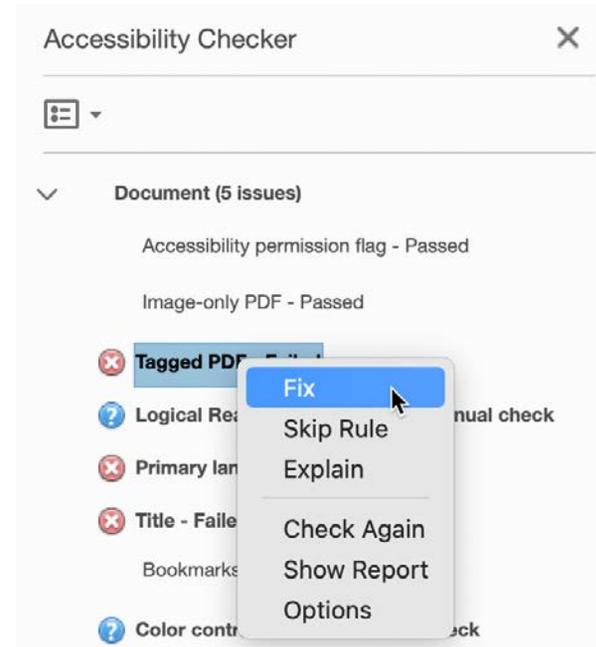


Running the Add unique ID to Note elements in tagging structure fix-up.

Error: Document is not maked as tagged

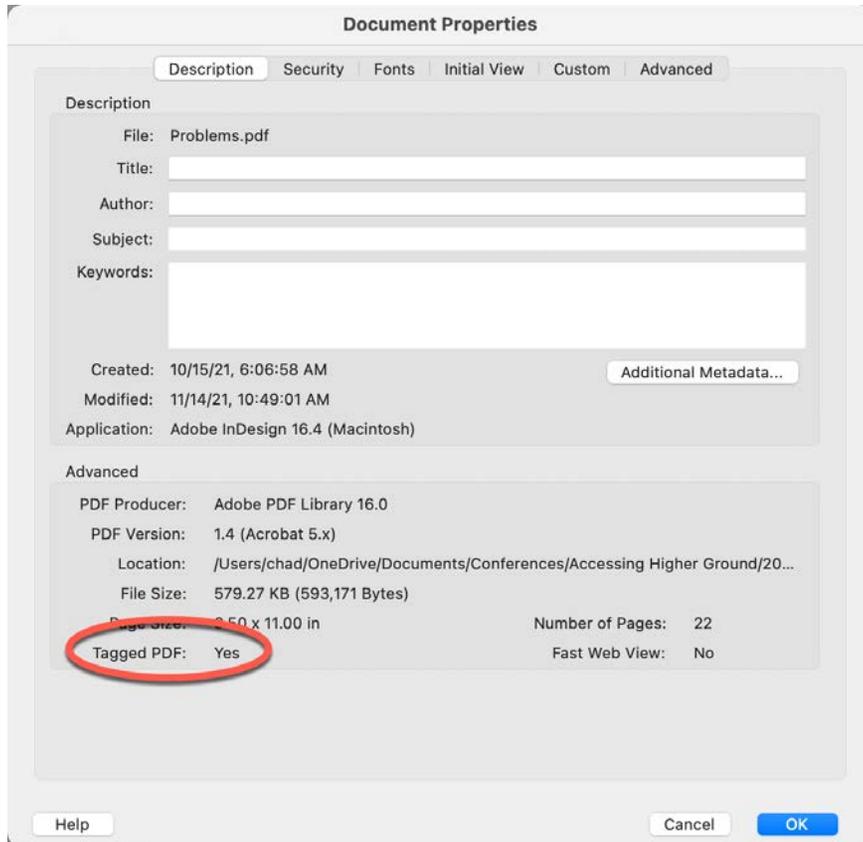
For PDF accessibility a flag needs to be added to the PDF file that makes assistive technology applications aware that the document is tagged and to look for the tags in a document. The “tagged” flag is strictly metadata and doesn’t appear anywhere visually in the document.

Fix: The easiest way to fix this error is to run the Acrobat Accessibility checker within Adobe Acrobat. In the resulting report, open the Document category and click on the Tagged PDF - Failed error. Right click on the error and choose fix. The error disappears and now shows as passing because the “tagged” flag has been applied to the document.



Fixing the Tagged PDF - Failed error in the Accessibility Checker report.

To see this flag applied in the document, choose File > Properties in Adobe Acrobat and click on the Description tab at the top of the Document Properties and in the Advanced section at the bottom of the dialog box, you'll see Tagged PDF: Yes.



Viewing the "tagged" flag applied in the Document Properties dialog box within Adobe Acrobat.