Ten common PDF accessibility errors with solutions

Table of contents

List of Figures

1. Why bother about accessible PDFs?

Each year, there is an increasing number of students with disabilities in higher education institution like the University of Ottawa. Most institutions have suddenly moved from traditional forms of teaching and learning ("chalk and talk") to online learning. Although presenting information online has a great potential for inclusion of people with disabilities, it equally has the same potential for their exclusion if care is not taken to ensure that electronic systems and documents are designed accessibly. Some people with disabilities rely on assistive technologies such as screen readers to be able to read information electronically. When this information or electronic systems are not designed with consideration of the needs of the users in mind, particularly those with disabilities, the result is exclusion. For instance, when a text book or document is scanned and produced as an image-only PDF, it is very difficult, if not impossible for screen readers to read the text in that image. Hence, a screen reader user will be excluded from obtaining the information presented in that format.

When presenting information on websites, it is therefore necessary to consider the needs of users with various types of disabilities. Fortunately, there are some guidelines on producing accessible websites and accessible documents such as PDFs. For more information on these, please read PDF Techniques for WCAG 2.0 and Web Content Accessibility Guidelines (WCAG) 2.0.

The importance of accessible PDFs cannot be overemphasized. Accessible PDFs are easy for people with disabilities to access either with assistive technologies such as screen readers for those with visual impairments or via the keyboard for those with mobility difficulties of their hands. By following existing guidelines and producing accessible documents, people with disabilities will be included and can obtain the information that is intended for them.

2. Common PDF accessibility errors and their solutions

Having reviewed over 1500 PDF documents at the University of Ottawa, this section presents some common errors encountered including how these can be resolved. This guide will be very helpful for those who want to improve on the accessibility of their PDFs.

2.1 PDF not tagged

When PDFs are tagged, the content is structurally divided into various sections. This helps give some order to the document. If the PDF is logically ordered, then screen reader users can be able to read the information without difficulties. Also, people who depend on the keyboard rather than the mouse to browse the PDF will be able to navigate the document with ease. Tagged PDFs provide the highest level of accessibility and when a document is not tagged, it often throws other accessibility errors when checked automatically either using Adobe Acrobat or other tools such as the PDF Accessibility Checker (PAC). On the contrary, when a PDF document is tagged, the accessibility is greatly improved.

To know if a PDF document is tagged or not, go to the File menu and select Properties to show the Document Properties window (Figure *1*).

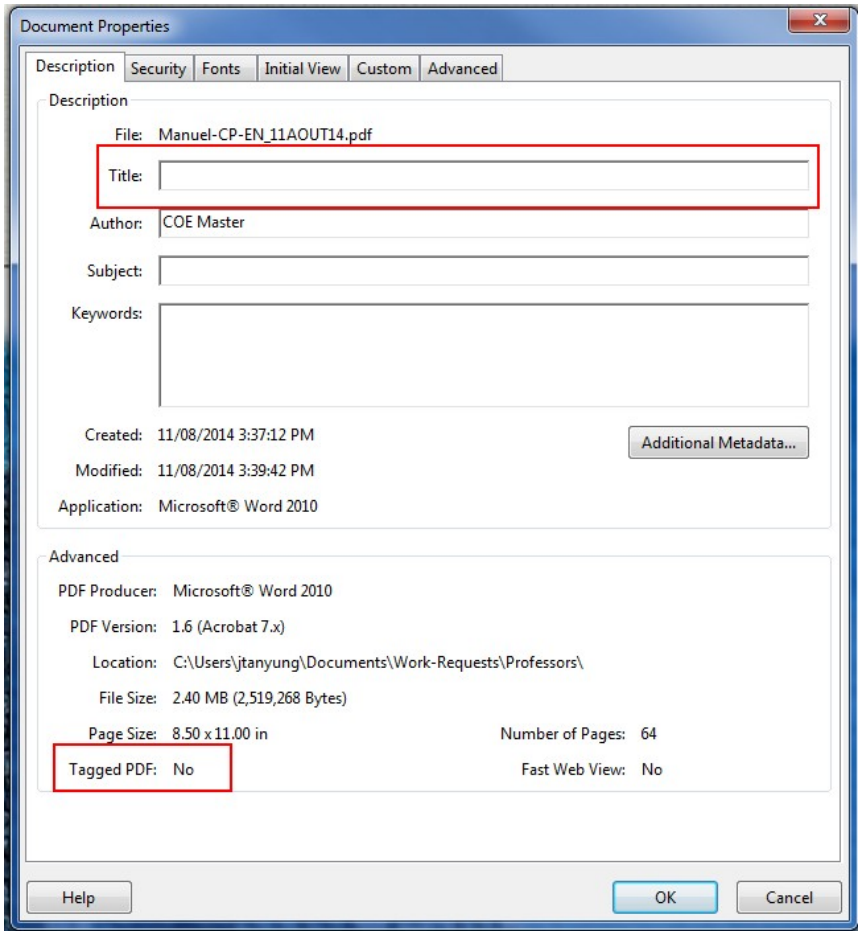

Figure 1: PDF Document Properties

Under the Advanced section of the Description tab, you can see the status of the PDF if it is tagged or not. In the case where the PDF is not tagged, you can tag it in Adobe Acrobat 9 Pro by going to the Advanced menu of the PDF document and selecting Add Tags to Document (Figure 2).

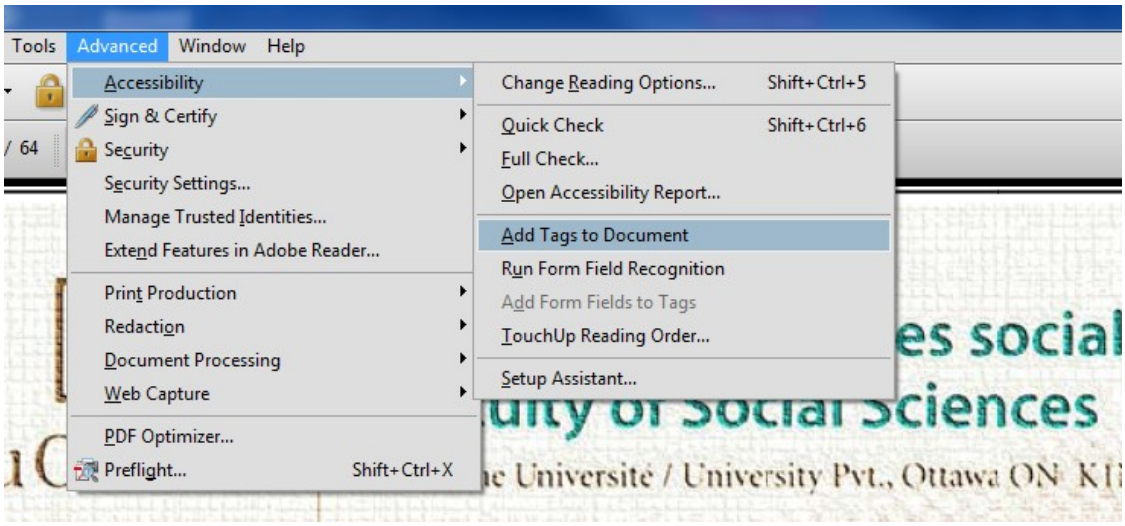

Figure 2: Adding tags to PDF documents with Adobe Acrobat 9 Pro

Once this is done, tags will be automatically added to the document, which will then increase the document`s accessibility as screen readers and other assistive technologies will be able to read the document. The tags added may not be very meaningful as they are automatically done but this helps improve the accessibility of the document.

After producing your PDF document, before you post it online or send to someone; please verify that the document is tagged. If this is not the case, please add tags to the document to improve its accessibility.

2.2 PDF not having a title

Even after producing a tagged PDF which has an improved accessibility, a PDF document without a title is still not fully accessible. This is because the title helps in identifying the document and without a meaningful document title; some users might not know they have the correct document.

Usually, if converting to PDF from other formats such as Word, the resultant PDF will contain the document title. If you forgot to enter the document title and then converted to PDF, you can still add a title by going through the document properties as shown in *Figure 1*. In the Description tab, please ensure that the title of the document is included. If not, then add a meaningful title that reflects the

contents of the document and in the language of the document. Please do not add an English title to a French document and vice-versa.

2.3 PDF title present but not meaningful

To help in identifying the document, the title needs to be meaningful and reflect the content of the document. Some authors convert into PDF from other formats and instead of having a meaningful title, the Title field contains the file path which is not meaningful. Some documents converted into PDF have titles such as "Slide1". A user will not understand what this is as the title is not meaningful. A good document title should contain a brief description of the content of the document. This will help the user understand what the document is. Avoid using file extensions in the document title such as .docx, .doc, .xls, .xlsx etc. After having produced your PDF, please ensure that the title is meaningful. Go through the document properties as in Figure *1* and then enter a meaningful document title in the language of the document.

2.4 PDF document language not defined

Defining the document language is important for accessibility purposes because a defined document language helps assistive technologies render text more accurately. Additionally, screen readers can load the correct pronunciation rules while visual browsers can display characters and scripts correctly. Also, media players can show captions correctly. As a result, users with disabilities are better able to understand the content.

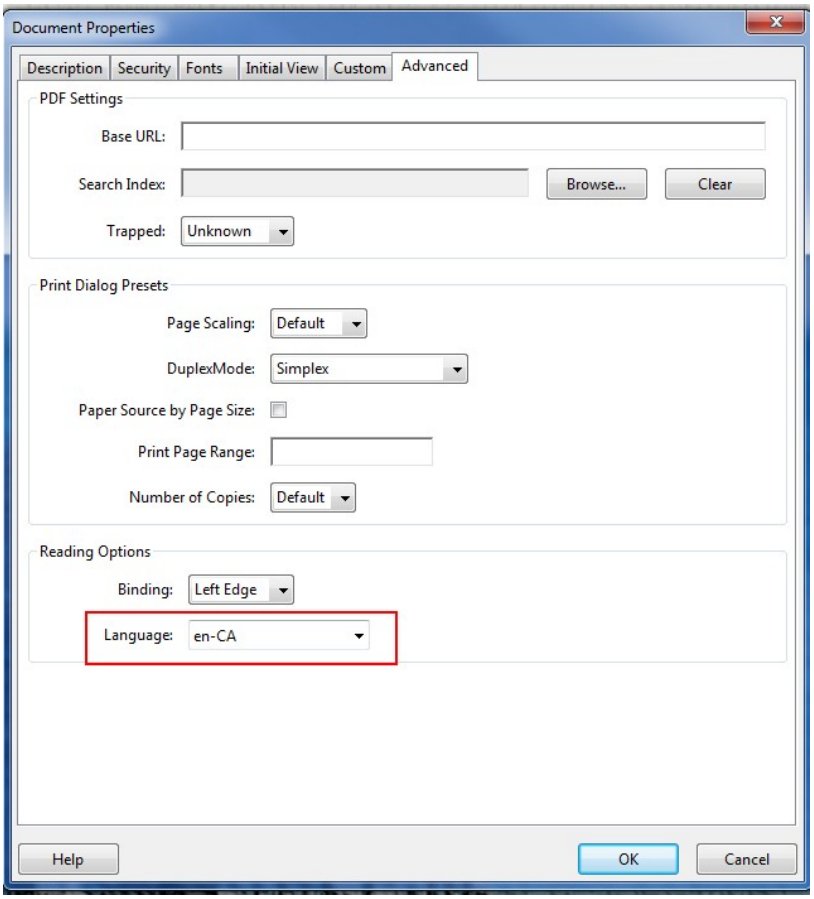

Figure 3: Defining the document language

To define the document language in Adobe Acrobat, open the document, go to File and select Properties, select Advanced tab and in the Reading Options field, select the default language as shown in (Figure 3).

2.5 PDF documents being bilingual

It is preferable to produce unilingual documents. The idea of producing bilingual documents particularly forms was probably translated from printed forms which were usually bilingual and users could then fill the forms in their language of choice. Whilst this idea is cost effective for the printed document, it poses accessibility challenges when in electronic format and users with assistive technologies need to access and fill them electronically. When a document is bilingual, assistive technology will not be able to pronounce some of the words correctly. If the language is not switched automatically, it is difficult for the user to keep trying to read a line of text in one language and then trying to switch over to the

other language to read the next line and so forth and this is time consuming and may be frustrating. This might result in the user abandoning the task.

There is even the issue of selecting the default language for the document. If it is bilingual (French and English), which language will be the default? It is thus preferable to produce a unilingual document and if necessary, indicate somewhere in the document that the same document is available in the other language.

2.6 Images not having alternative texts

To comply with WCAG 2.0, each image must have a text alternative that serves the equivalent purpose. This is very useful for people with visual impairments. When an alternative text is not available, a screen reader user will not be able to understand what the image represents. To facilitate the task of adding alternative texts to all images, ensure that before converting into PDF from Word, all images have alternative texts while background images are labelled as decorative.

Images that are purely decorative or background images do not need alternative texts. In this case, they can be marked as decorative or "artifacts" so that assistive technology will ignore them. When marked as such, the document will not present accessibility errors resulting from missing alternative texts.

To add an alternative text to an image in a PDF document with Adobe Acrobat 9 Pro, go to Tools, select Advanced Editing and then TouchUp Object Tool. Right click on the image and then select Properties. In the Tag tab, enter the alternative text for the image in the Alternative Text field (Figure 4).

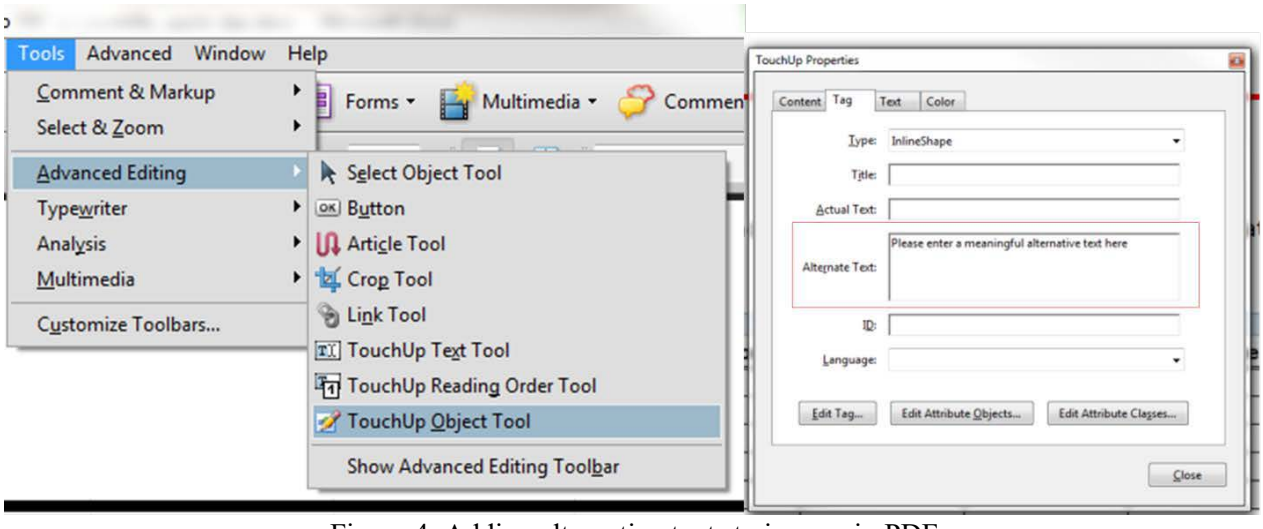

Figure 4: Adding alternative texts to images in PDF

Decorative elements in the PDFs such as decorative images, headers, footers, etc. can be marked as "artifacts" so that assistive technologies can ignore them. To do this in Adobe Acrobat 9 Pro, select the image you wish to designate as Background then open the TouchUp Reading Order tool by selecting Tools in the main menu, Advanced Editing and then TouchUp Reading Order Tool (Figure 5).

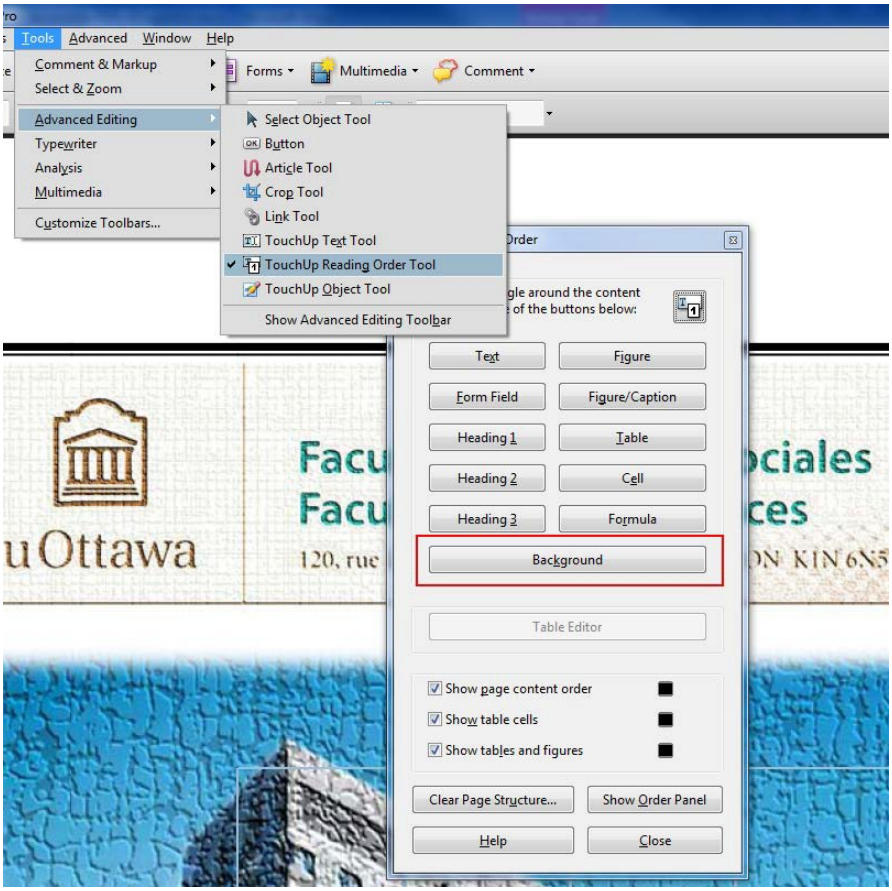

Figure 5: Designating images as "artifacts" or "background" in Adobe Acrobat 9 Pro

In the TouchUp Reading Order tool, click the Background button in order to remove the selected image from the tag structure. By doing this, assistive technology will ignore the image when reading the document and this will not be flagged as an accessibility error when checking the document for accessibility.

2.7 Long PDF documents not having bookmarks

Bookmarks help users navigate the PDF document and find information. This is very helpful for people with cognitive disabilities. When bookmarks are available, users can readily move to specific

sections of the document by clicking on the bookmarks. Hence, it is necessary to not only add bookmarks but to add meaningful bookmarks which represent the various sections of the document.

If creating the document in Word before converting into PDF, please ensure that you create a table of contents. Additionally, use the heading styles found in Word and ensure that the headings are correctly applied. So, a Heading 2 should follow a Heading 1 and so forth. When this is properly done and the final document is converted into PDF, the resultant PDF will contain bookmarks. This helps improve the accessibility of the PDF.

Whilst bookmarks are generally recommended for long PDF documents, adding bookmarks to a document with even a single page helps improve its accessibility and so authors are encouraged to add bookmarks to all PDFs they produce.

To add bookmarks to an existing PDF document, in the Bookmarks panel, choose the options menu and then select New Bookmark from Structure (*Figure 6*). Now select the elements you want to specify as tagged Bookmarks from the Structure Elements dialog box. The bookmarks will be presented under an Untitled section which can then be renamed by accessing the context menu and selecting Rename.

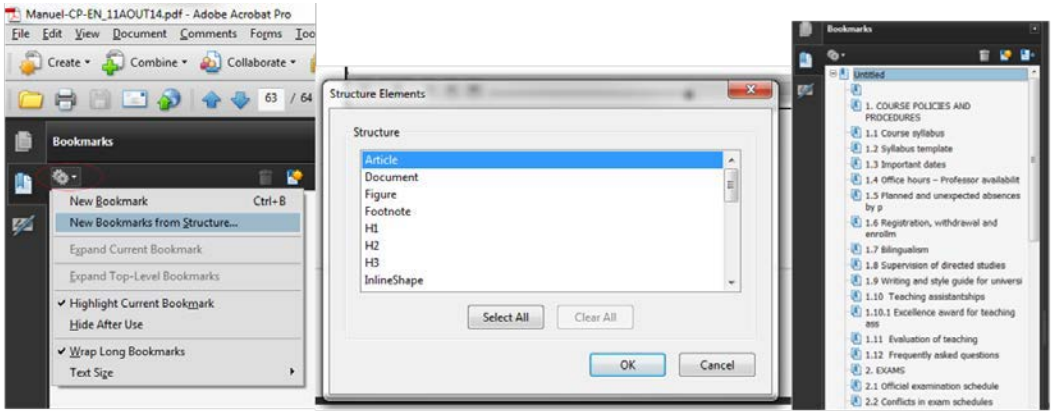

Figure 6: Adding bookmarks in Adobe Acrobat 9 Pro

2.8 Tab order not following the document structure

Tabs are used to navigate a PDF document. This is very useful for keyboard and assistive technology users who depend on the keyboard and not the mouse to navigate the document. It is good to ensure that users can navigate through content in a logical order that is consistent with the meaning of the content.

To set the tab order to follow the document structure in Adobe Acrobat, go to the Pages panel, right click on a page thumbnail and then select Page Properties. In the Page Properties dialog box, select the Tab Order tab and then choose Use Document Structure as shown in Figure 7. This ensures that the tab order follows the document structure. Hence an accessibility error related to this tab order will not show up in accessibility reports.

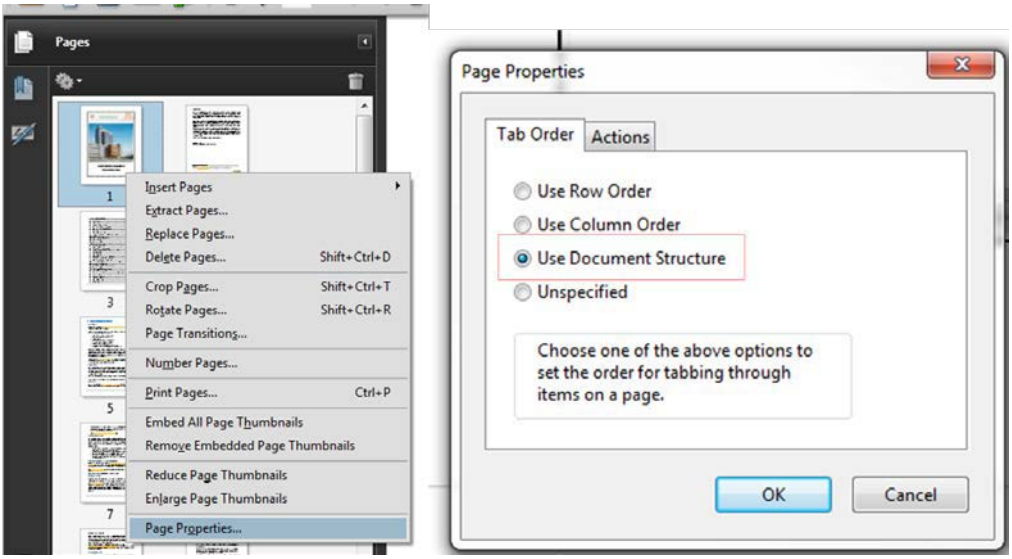

Figure 7: Setting the Tab order to follow the document structure in Acrobat

2.9 Table headers not defined

Defined table headers are required for accessibility. When the tables are well marked up, they will be recognized by assistive technology. To correctly do this in Adobe Acrobat 9 Pro, go to the View menu and then select Navigation Panel, then select Tags. On the Tags tab, open the table row that contains the header cells as shown in Figure 8. Select the first data cell and right click then select Properties. In the Tag tab of the Properties dialog, under Type, select Table Header Cell and then repeat this for all header cells in the first table row.

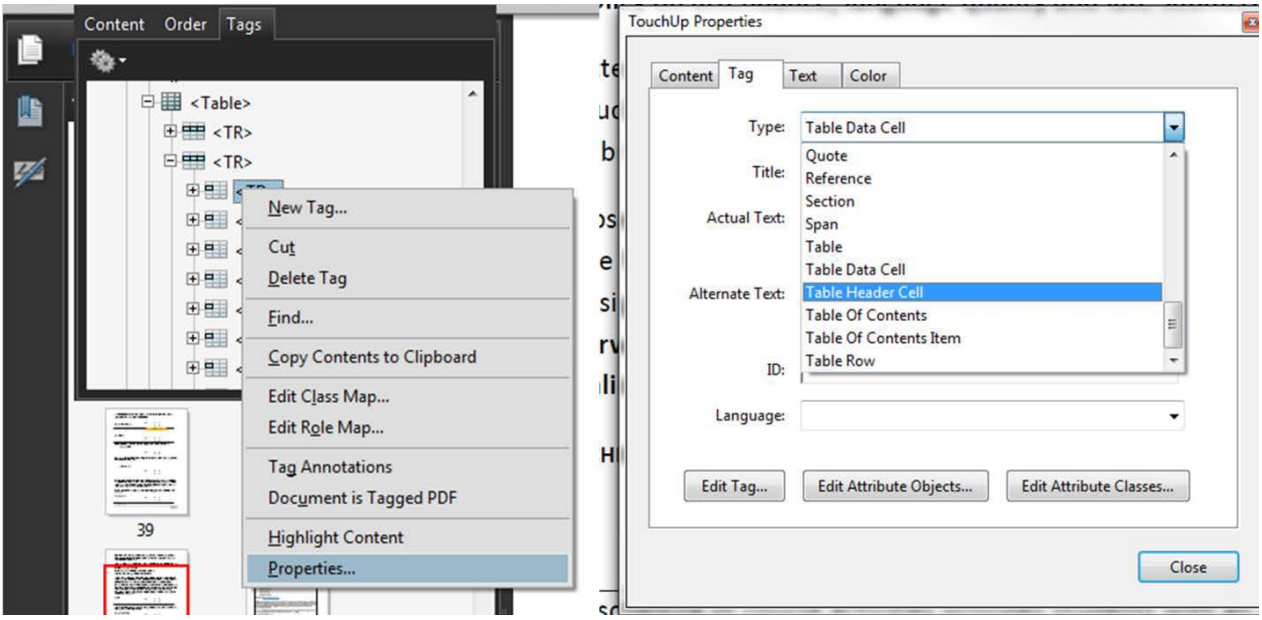

Figure 8: Defining table header cells in Adobe Acrobat 9 Pro

2.10 Image-only PDFs

When a PDF is made up of an image only, it is inaccessible as screen reader users will not be able to read its content. To ensure that such visually rendered text is presented in a manner that can be perceived without the visual presentation interfering with its readability, this needs to be converted into a text-based PDF using optical character recognition (OCR) which is a feature of Acrobat.

To accomplish this, go to the Document menu, from the OCR Text Recognition sub menu, select recognize text using OCR then select which pages you want to convert into text and then select OK (Figure 9).

3. Checking the accessibility of PDF documents

When documents have been converted into PDFs, it is important to know how to verify that they are accessible. Even when the author has been very careful in ensuring that the document is accessible in the original format (Word, PowerPoint, etc.) there could still be some accessibility errors that need to be checked.

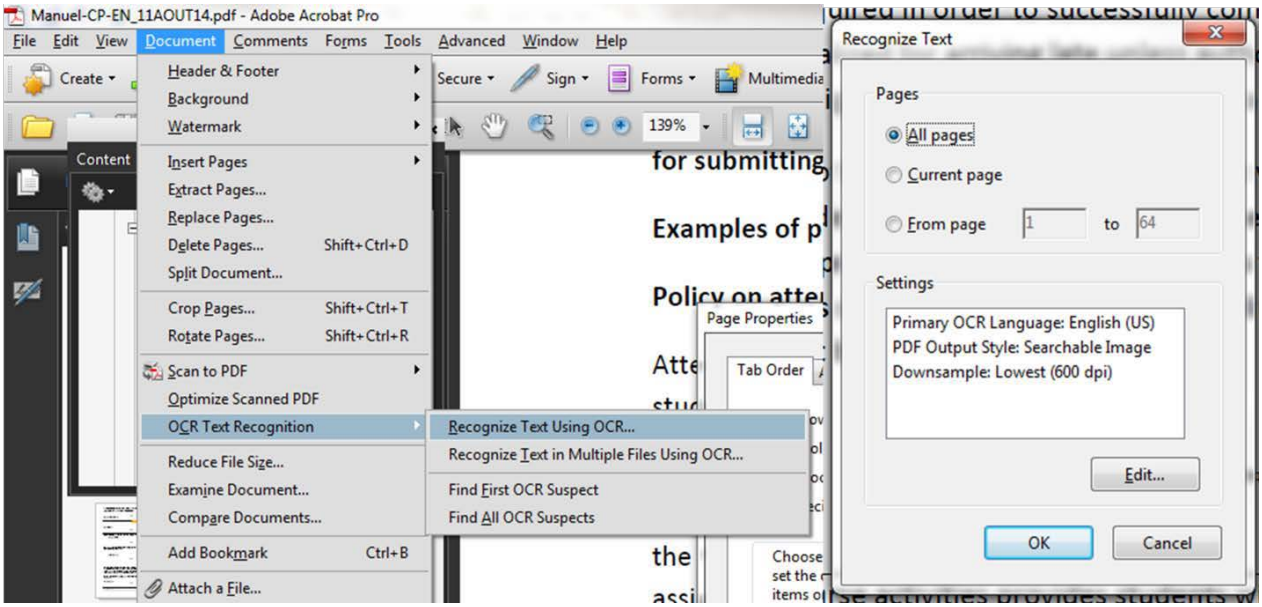

Figure 9: OCR text recognition in Adobe Acrobat 9 Pro

There are several tools that can help in checking the accessibility of PDFs but we will recommend using Acrobat accessibility checker in addition to the PDF Accessibility Checker (PAC). Acrobat`s accessibility checker is inbuilt and Acrobat 9 Pro and versions above all have this tool. PAC can be downloaded and used to check accessibility. Whilst these tools help in checking the accessibility and finding some errors which can be fixed, they cannot find all the errors in the document. For instance, an author might just enter some numbers in the title field and it will be recognized as having a document title whereas a human expert will be able to determine that it is not a meaningful document title and does not reflect the content of the document. Also, the file name can be entered in the title field and will be recognized as a title by the automated tools whereas a human expert will flag this. Therefore, using a combination of methods helps find accessibility and enable the author to improve the accessibility of the document.

3.1 Using PDF accessibility checker (PAC)

PAC has two versions (PAC 1.3 and PAC 2). Whilst PAC 2 is the latest version and much improved, PAC 1.3 still functions very well and will flag many accessibility errors. Both can enable the user to click and view the errors in detail in order to determine how to resolve it.

The criteria used by PAC 1.3 are shown in Figure *10*.

Figure 10: Criteria used by PAC 1.3

3.1.1 Using PAC 1.3

To use PAC 1.3, first download PAC 1.3 and unzip it. From the folder containing the PAC application, double click the file PAC.exe and then select Run in order to run PAC 1.3. Then click Browse and then select the PDF file you wish to check as shown Figure *11*. Once the file has been selected, click Start Check to check the file. You will then obtain the results also shown in Figure *11*.

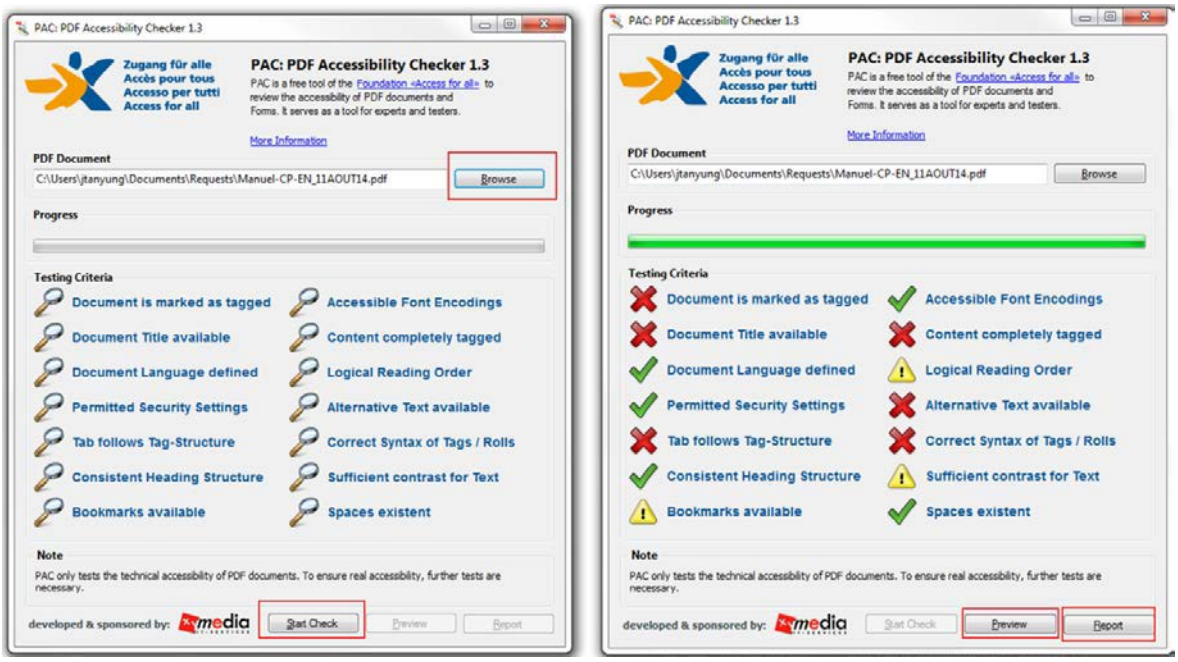

Figure 11: Checking a PDF file with PAC 1.3

Once in the Report screen. You can click on preview to have an overview of the structure of the document. From the Report screen, clicking on Report will reveal details of the errors that are in the PDF. There are also links where you can click to find out more about the error as shown in Figure *12*.

Figure 12: Detailed PDF accessibility audit in PAC 1.3

3.1.2 Using PAC 2.0

To use PAC 2.0, first download PAC 2.0 and unzip it. From the unzipped folder containing the PAC 2 application, double click the file PDF Accessibility Checker 2.0.exe and then select Run in order to run PAC 2. Then click Browse and then select the PDF file you wish to check as shown in Figure *13*. Once the file has been selected, click Start to check the file. You will then obtain the results also shown in Figure **13**.

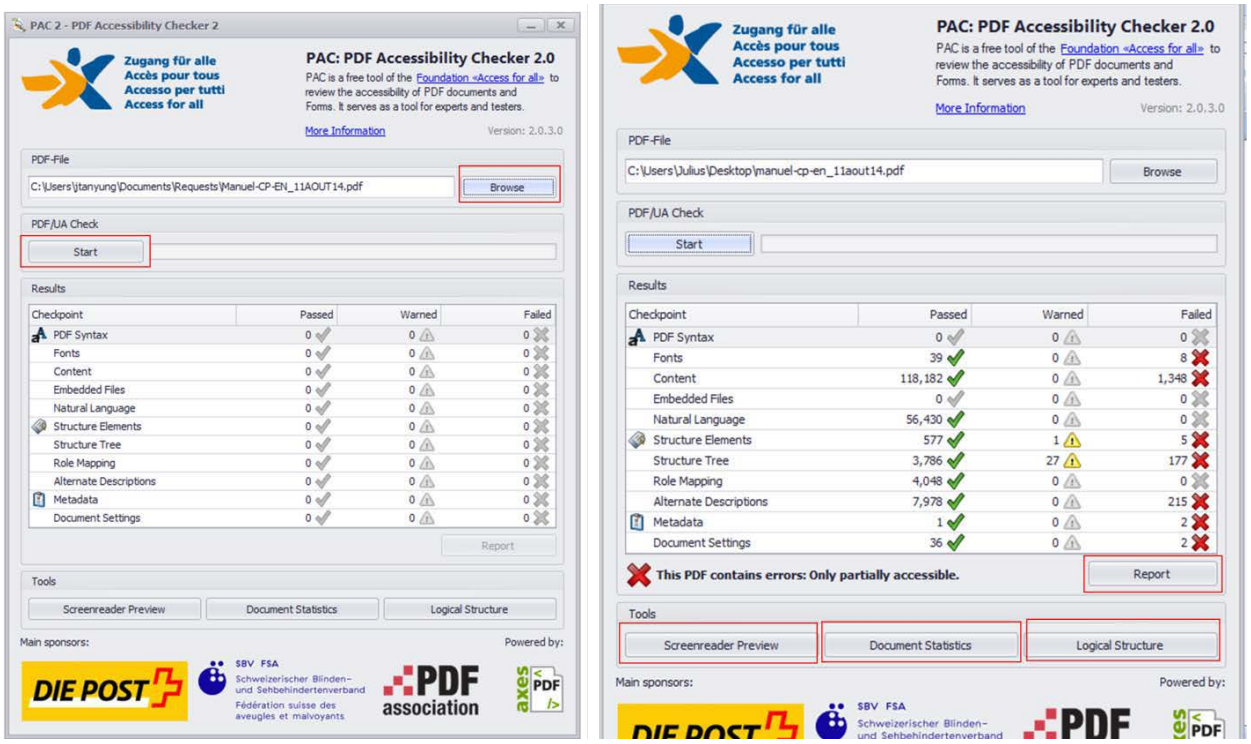

Figure 13: PDF accessibility audit with PAC 2.0

PAC 2.0 gives the user the option to view the report as a tree view by clicking Report. The tree view is shown in Figure *14*.

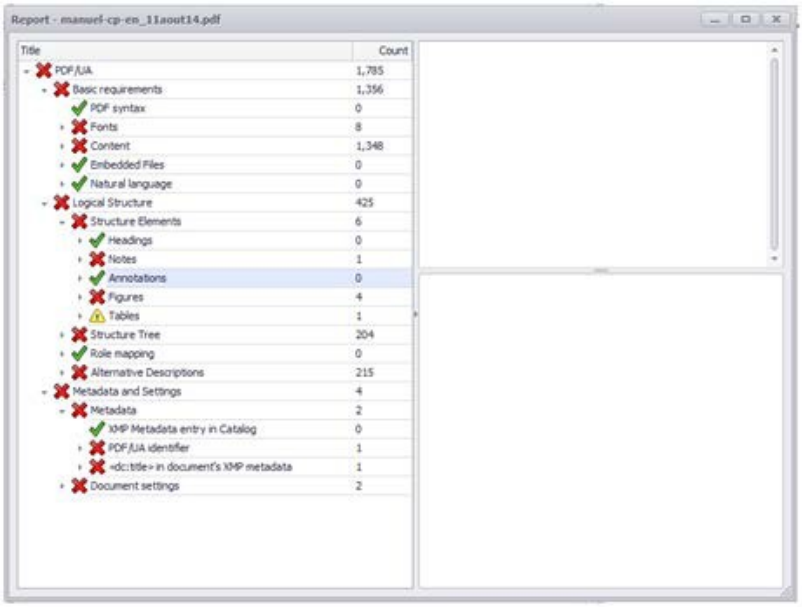

Figure 14: Tree view of PDF accessibility audit report with PAC 2.0

PAC 2.0 has some important tools:

3.1.2.1 The Screenreader Preview tool

This tool enables you to view the document layout based on how a screen reader will view it. An example of the screen reader preview is shown in Figure *15*.

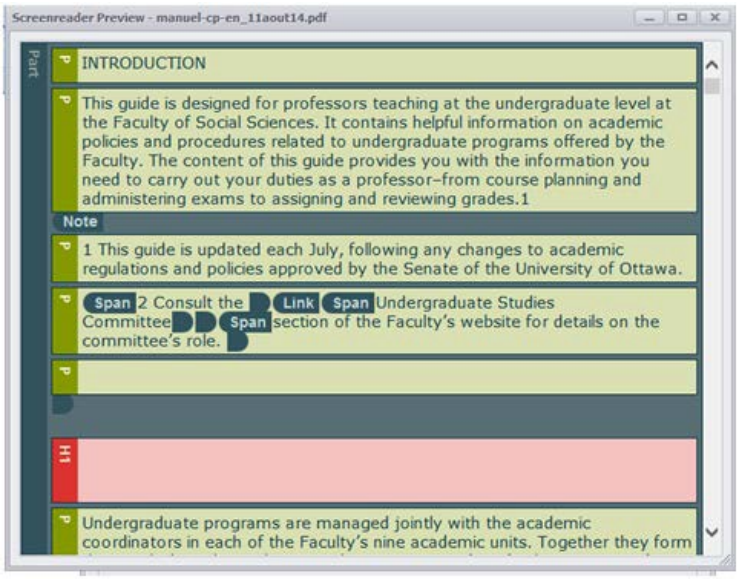

Figure 15: Screen reader preview in PAC 2.0

This preview is important as it helps you determine if your document's reading order is logical and well structured.

3.1.2.2 The Document Statistics tool

This tool gives statistics on the structural elements (e.g. figures, headings, tables, etc.) and annotations (e.g. links) found in the document. Please see Figure *16* as an example.

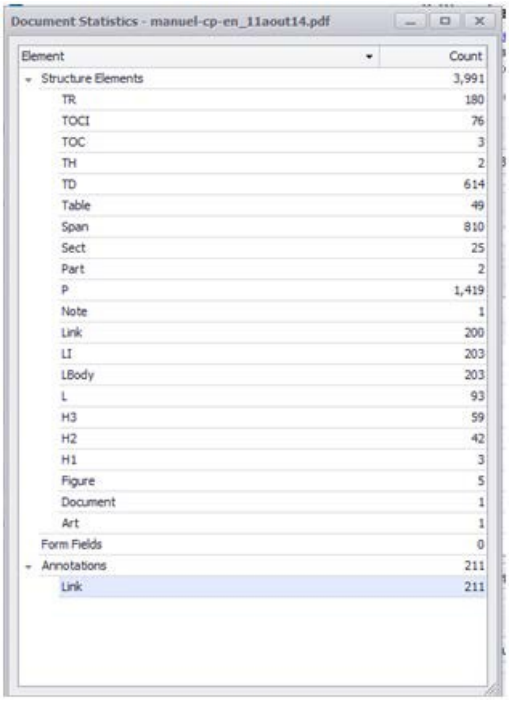

Figure 16: Document Statistics view in PAC 2.0

3.1.2.2 The Logical Structure tool

With this tool, you can preview the structure of the document with the options to view the document properties, a page view of the document (Figure *17*) and a role map.

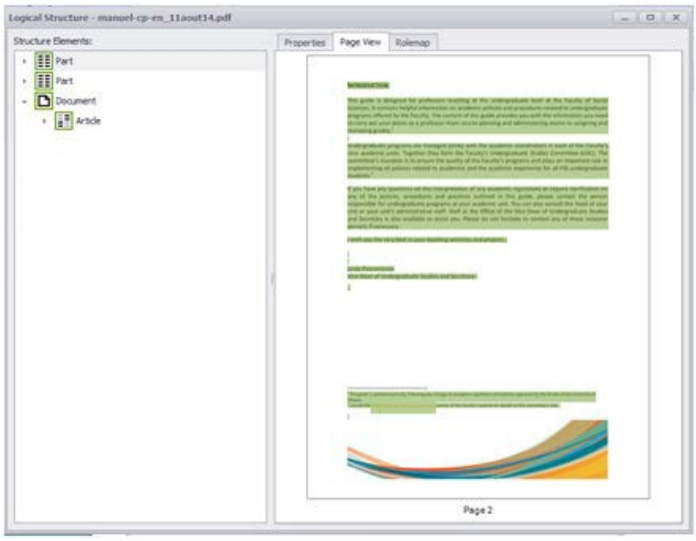

Figure 17: The Logical Structure view in PAC 2.0

3.1.3 Common PDF/UA compliance errors

Sometimes when checking with PAC 2.0, you may come across some errors that Adobe Acrobat will not find. Some of these errors are due to compliance with PDF/UA. Some of the errors as identified in the Adobe Forum online and their solutions are as follows:

3.1.3.1 PDF/UA identifier missing

Sometimes, some documents can pass the Adobe Acrobat checker as well as PAC1.3 but could have an error that says *PDF/UA identifier missing*. To fix this error, create an xmp file that includes the required snippet of metadata or simply download one then do the following once the document is open:

- Go to the file menu and select Properties.
- In the description tab, click on Additional Metadata
- Click on Advanced and then on Append
- Select the XMP file then click OK.

3.1.3.2 Font not embedded

When you get this error, in order to fix the following:

- Go to Tools and select Print Production
- Select Preflight and then PDF fixups

- Select Embed fonts, Click on Analyze and fix. It may be that the PDF will have to be saved again. Save it.

In the event that this fix does not work because the fonts cannot be embedded, go back to the original document from which the PDF was created and then use a different font.

3.1.3.3 Alternative description missing for an annotation

This usually happens when the link is not clear enough. To fix this error, add alternative text to the link tags. In order to add the alternative text, please do the following;

- In the tag tree, select the <link> tag for the link, and select Properties from the options menu.
- In the TouchUp Properties dialog box, select the Tag Tab.
- Type alternate text for the link, and click close.

3.1.3.4 Link annotation is not nested inside a link structure element

This might occur when the link annotation is nested inside a <reference> element instead of <link> element for example or when a <link> element exists but the link annotation is not nested within it. If the link annotation is found within a <reference> element instead, this can be corrected in Adobe Acrobat XI Pro by doing the following:

- Expand the Tags tree and right click on the <reference> element.
- Select Properties.
- In the Tag tab, Select Link as the Tag type.
- Click Close and then save the changes in your PDF document.

3.1.3.5 Sect structure element used as root element

This may be a warning for the fact that there exist some Section structure elements in the document that are not used semantically correctly. For instance, instead of starting the document structure as <Document>, <Section> is used. If the PDF was converted from PowerPoint, you might have many sections tagged as <slide>. Follow the same procedure to change all to <Document>.

To fix this warning or error, do the following:

- In the Tags tree, right click on the <Sect> element that has been used as a root element.
- Select Properties

- In the Tag tab, under Type, select Document.
- Click Close
- Save the PDF document.

3.1.3.6 Figure element on a single page with no bounding box

Tag the figure as artifact if it is not a figure that needs to be interpreted by a screen reader. To do this, find the element in the Tag tree and then right click on it and select "Change Tag to Artifact".

3.1.3.7 Alternative text missing for Figure structure element.

Tag the figure as artifact if it is not a figure that needs to be interpreted by a screen reader. To do this, find the element in the Tag tree and then right click on it and select "Change Tag to Artifact". If you need to add an alternative text, right click on the element in the tag tree and select Properties. In the Tag tab, enter the alternative text for the image in the Alternate Text field. Click Close.

3.2 Using Adobe Acrobat accessibility checker

We will not spend much time discussing Adobe Acrobat as there is a very useful guide titled: Create and verify PDF accessibility (Acrobat Pro). Once your PDF file is ready, to check its accessibility in Adobe Acrobat, open the document using Acrobat. You can choose to run a quick check or a full check. A full check will show you the errors in the PDF document with links that enable you to view details about the error and how to fix it. To Run a full check in Adobe Acrobat 9 Pro, go to the Advanced menu then select Full Check from the Accessibility sub menu as shown in Figure *18*.

Figure 18: Running an accessibility Full Check with Acrobat 9 Pro

An Accessibility Full Check dialog will appear giving you the option to run the accessibility based on Adobe PDF criteria, Section 508, WCAG 1.0 or WCAG 2.0. Once you select your option and click Start Checking, the results will be displayed to the left of the document under the Accessibility Report

section as shown in Figure *19*. In the results section, there are tips on how to repair the PDF document.

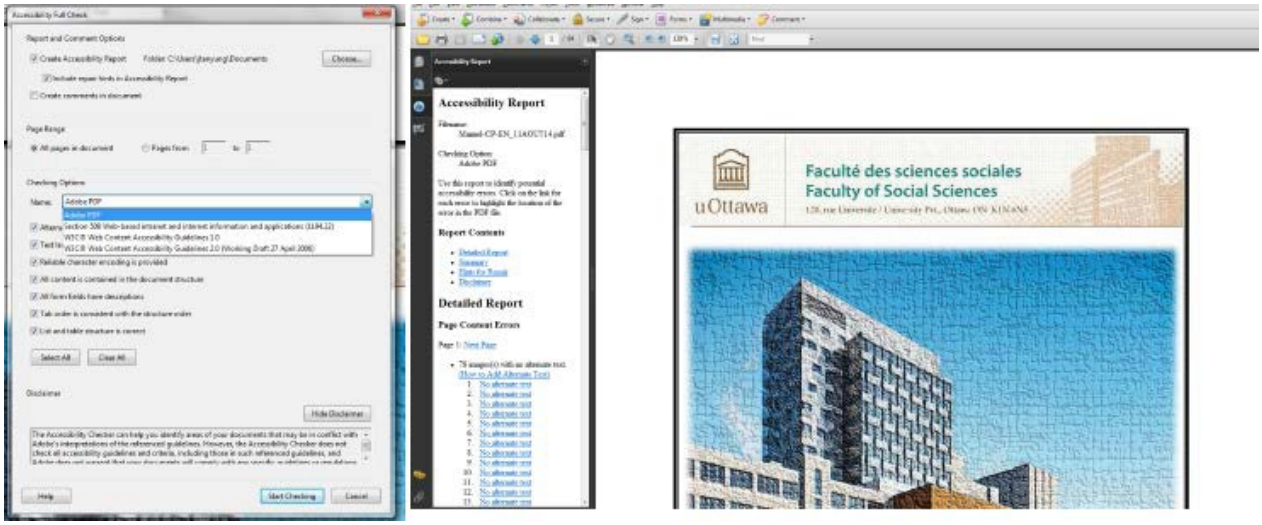

Figure 19: Results of accessibility Full Check with Acrobat 9 Pro

3.3 Using human expertise

As earlier discussed, even when automated tools like PAC and Adobe Acrobat are used to verify the accessibility of the PDF, we still need a human expert to find the errors that could not be detected by the automated tools. In checking the PDF, please ensure that the document title is meaningful and reflects the content of the document and that the title is in the language of the document. Ensure that there is sufficient color contrast.

You could also use an assistive technology or have someone who uses one read the document with the assistive technology and report any accessibility errors they find. This way, you will be sure to find most of the errors and then fix them.

4. Useful PDF accessibility resources

This guide cannot cover all that one needs to know in order to make a PDF document fully accessible. Fortunately, there are several free resources that are available online to help you improve the accessibility of your PDF document. Amongst the numerous resources which we cannot all list are:

1. Web Content Accessibility Guidelines (WCAG) 2.0
2. PDF Techniques for WCAG 2.0
3. PDF Accessibility- Resource from WebAIM
4. Adobe Acrobat Accessibility Training Resources
5. Tagged PDF: 508 PDF Help Center
6. An assessment of the accessibility of PDF versions of selected journal articles published in a WCAG 2.0 era (2014–2018)
7. The Portable Document Format (PDF) accessibility practice of four journal publishers