

VAN Briefingxtra

The voice of the Voluntary Arts

Understanding and creating accessible PDFs

This briefing is an introduction to the issues faced by disabled users in accessing and using PDF documents. It contains practical information about creating accessible PDFs using two common word processing packages - OpenOffice Writer and Microsoft Word. It also provides some basic information about professional packages and how to commission accessible PDFs from designers. It is not intended for use by designers or those using professional desktop publishing tools.

A. Introduction

Portable Document Format (PDF) files are now common and used on a daily basis. Increasingly, you'll find the information is not available in any other format. The user manuals of many new machines, like printers and iPods, for example, are often only available in PDF format. Additionally, as more and more services such as insurance and banking move online, customers will increasingly be required to use PDF documents.

The aim of the PDF format is to make documents portable. This allows you to create a document in one application, such as Word, OpenOffice, InDesign or GoogleDocs, and 'export' it as a PDF so that it can be read by people who don't have access to the software you used to create it.

Since PDFs are now so ubiquitous, you would imagine that the ability of people with disabilities to use them would be a given. Unfortunately this is not the case. Whilst it is fairly simple to create PDF documents which can, in the main, be

used by people with disabilities, the vast majority of PDFs are not created with accessibility in mind. Their level of accessibility is reliant on the default settings of the authoring software (Word, OpenOffice etc).

For example, consider a PDF of a company's annual report. The report presents the company's figures for the year as a series of pie charts and graphs, which take the form of JPEG images. To the sighted user this way of presenting information doesn't cause any problems - but to the non-sighted user, unless the images have an 'alternative text description' (*more about this later*) attached to it, the user may be aware of the presence of an image, but will have no more information than that.

***A note about terminology:** accessible PDFs are often referred to as 'tagged PDFs'. These terms are broadly interchangeable and both are used in this briefing.*

B. The debate about PDF accessibility

It is not the intention of this briefing to explain in detail the debate that surrounds PDF accessibility, but, in brief, most experts in the field now agree that in the majority of instances PDFs, if correctly authored, do not exclude disabled people from being able to access their contents. However, some experts assert that, in many cases, the information should also be provided in an alternative format. If the PDF is intended for use online, the most appropriate alternative in most instances would be as an HTML (Hypertext Markup Language) document. For standalone documents an RTF (Rich Text Format) would generally be the best alternative.

However, should it not be possible to provide an alternative - for budgetary reasons or because the necessary skills or software are not available to you - ensuring that your PDFs are as accessible as possible would be the next best alternative.

C. Why make PDFs accessible?

I. It benefits disabled users ...

Disabled users vary enormously in the way they read PDF documents. Some users may read the document by magnifying the document on screen, others may use assistive devices such as a screen or braille reader, whilst others may have a physical disability and therefore be unable to use a mouse. Taking the time to include your disabled users by ensuring that your PDFs are tagged and as accessible as possible will mean that these users are not excluded from your services or unable to make use of your information.

For more information on assistive technologies go to www.papergecko.co.uk/news/introduction.html.

2. ... and everyone else benefits too

As with web accessibility, ensuring your PDF documents are accessible offers several additional benefits. These include:

- **Improved search engine visibility** - if your PDF documents are intended for use online, making them accessible will improve their visibility to search engines such as Google (since search engines to some extent work in a similar way to assistive devices).
- **Better navigability** - making your PDFs accessible will often involve creating bookmarks or other aids that make navigation easier, thereby also making your PDFs easier to use for everyone.
- **Improved structure** - accessible documents are, by their nature, well structured, a factor which will improve all your users' experience of your documents.
- **More portable** - when you create an accessible PDF, you add XML (eXtensible Markup Language) information (tags) to your document. XML documents are extremely portable, which means that it will be much easier to reuse the contents of your documents in another form.

D. How do we create accessible PDFs?

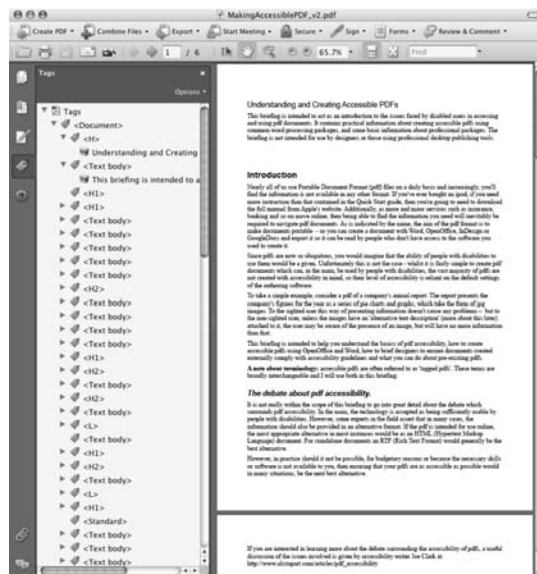
I. The software

PDFs can be created with a huge number of different software packages. Indeed, if you use a Mac (with a recent version of OSX) then PDF creation forms part of the operating system and is available via the print dialogue screen. However, software packages that can create **accessible** PDFs are much more limited.

The main ones are:

- **Microsoft Office** (Windows only)
- **OpenOffice**
- **Adobe InDesign**
- **Adobe Acrobat and Acrobat Professional** (but not Acrobat Reader). The Acrobat family of applications has a special status in that they should be able to make accessible PDFs from virtually any application that can print. Additionally, Acrobat Professional is able to convert existing, untagged PDFs into tagged, accessible versions.

2. Structure is king



Screenshot of Acrobat Professional, showing tag structure (left hand pane) in an accessible PDF.

For anyone familiar with HTML (Hypertext Markup Language - the language used to create web pages) the structure of an accessible PDF should be quite familiar. PDFs are mainly made up of text, headings, lists, graphics and images and employ 'tags' in a very similar way to that used in HTML (even down to enclosing tags in < > symbols). And, as in HTML, the accessibility of your documents hinges largely on creating documents which employ a structure which is understandable when read by another device.

The key to making accessible PDFs is the use of a logical, semantic structure. In practice this means re-thinking the way you use structural elements like headings, body text, lists, tables and so on and using them in a way that assistive devices can interpret them. For example, you may be in the habit of creating headings in Word by increasing font size or by underlining, rather than by using the heading options available in the Styles palette. The former approach is perfectly

appropriate for sighted users, who can see that you are emphasising certain parts of your text. But increasing the font size does not affect the structure of your document at all, so users of assistive devices are offered no indication that certain parts of your text are to be read as headings. Using the Styles palette would affect the document's structure, as well as its appearance, so is much the best option for serving all your potential readers.

The main changes you will need to make are:

- **Headings** - use the heading structures built into the software package, ensuring that the hierarchical structure is correct. If you are using Word or OpenOffice this will have the dual benefit of creating a sensible system of bookmarks (based on your headings) and should also enable you to quickly create and amend tables of contents.
- **Lists** - if you are creating lists (bulleted or numbered), use the list functions provided by the software. Don't use bullets just for emphasis (instead of bold for example). Use numbered lists where appropriate (to indicate a series of steps for example).
- **Images** - all images need to be marked up with 'alternative text' (see Section E1 for more information).
- **Tables** - use the table functions provided by the software. Both OpenOffice and Word have the ability to indicate which table cells include column headings.
- **Tables of contents** - where documents are of sufficient length, consider providing a table of contents to assist readers in navigating the document (both as PDFs and in their native formats). Both Word and OpenOffice include an automated table of contents feature.
- **Bookmarks** - again, providing users with additional help to navigate the content will assist both disabled and non-disabled users of your document. Both Word and OpenOffice are able to create automated bookmarks using your heading structure.

E. Using OpenOffice Writer and Word

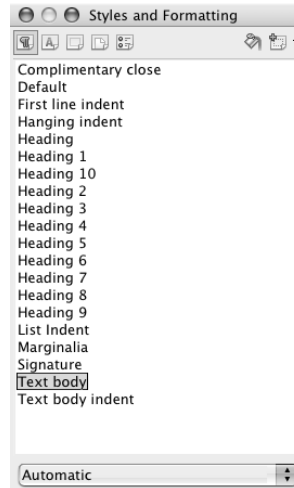
If you have access to Adobe Acrobat or Acrobat Professional you will be able to create accessible PDFs from virtually any software which is able to print. But the packages which are most often used to create text documents are Microsoft Word and its open source equivalent, OpenOffice Writer.

I. OpenOffice Writer

OpenOffice (www.openoffice.org) is a free, open source alternative to the Microsoft Office suite which is available for

Windows, Mac OS and Linux. This briefing looks first at creating accessible PDFs in the word processing component of OpenOffice (Writer) before discussing Microsoft Word, since Word's ability to create accessible PDFs is not consistent across the different versions and platforms. OpenOffice, on the other hand, is consistent across platforms and has supported the creation of tagged PDF documents since version 2.0.

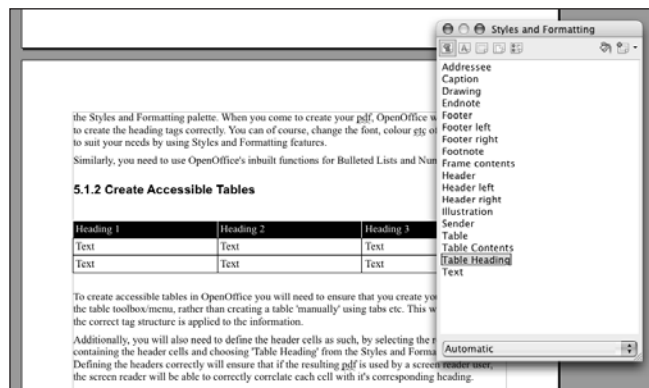
Use a logical structure



OpenOffice Writer's Styles and Formatting palette.

As has already been discussed, the key to creating accessible PDFs lies in using a logical structure in your document. You also need to use the software's built-in structural elements rather than creating your own. So if you are creating headings, you will need to use the heading options (heading 1, heading 2, etc) available in the Styles and Formatting palette. When you come to create your PDF, OpenOffice Writer will then be able to create the heading tags correctly. You can also change the font, colour etc of these headings to suit your needs by using Styles and Formatting features. Similarly, you need to use OpenOffice's inbuilt functions for bulleted lists, numbered lists, tables and so on.

Create accessible tables

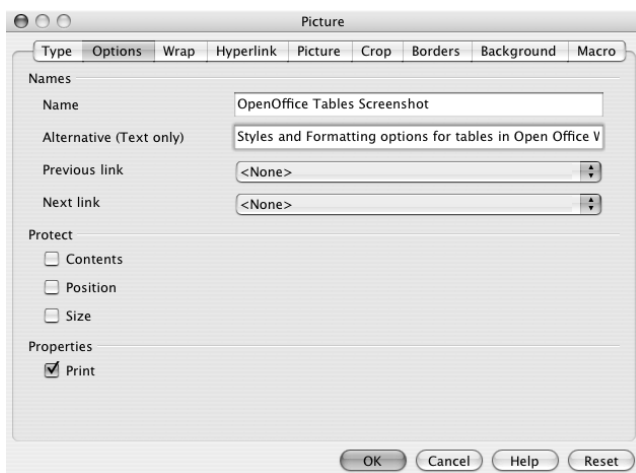


Formatting tables in OpenOffice Writer.

To create accessible tables in OpenOffice Writer you will need to ensure that you create your table using the table toolbox/menu, rather than creating a table 'manually' using tabs. This will ensure that the correct tag structure is applied to the information.

Additionally, you will also need to define the header cells as such, by selecting the row(s) containing the header cells and choosing 'Table Heading' from the Styles and Formatting menu. Defining the headers correctly will ensure that if the resulting PDF is used by a screen reader user, the screen reader will be able to correctly correlate each cell with its corresponding heading.

Give all pictures an alternative description



Adding alt text to images in OpenOffice.

Alternative text is intended to supply the non-sighted user with an equivalent to the image. In most cases this will be a description of the content of the image, its meaning or what it is intended to communicate. For example, if the image is of a pie chart, the 'alt' text 'Daily Internet usage: 10 per cent 12am to 8am; 30 per cent 8am to 12pm; 60 per cent 12pm to 12am' would be far more useful than 'Usage pie chart'.

So, if you are using images in your document, you will need to provide an alternative text description for each picture which describes the image's content for non-sighted users. To do this:

- Right click (or ctrl-click on a Mac) on the picture and choose Picture from the list.
- In the dialogue box that opens, click the Options tab and then enter your alternative text in the box marked 'Alternative (text only)'.
- When you have finished, click OK.
- Repeat this process for every picture in your document.

How to create appropriate alternative text

You don't usually need to state that it's a picture - e.g. if the image is of John Smith in the potting shed, the alt text would read 'John Smith in the potting shed' rather than 'a photo of John Smith in the potting shed' - the fact that it's an image is implied in the PDF tag. The exceptions to this are when the type of image presented is integral to its meaning. For example, if the document is an artist's catalogue then whether the image is a sculpture, print, or painting is likely to be important information.

If the image is a logo, or an image containing predominantly text (a 'Sale now on!' graphic for example), the alternative text should include exactly the same text as in the image. If the image is a logo, it is standard practice to include the word logo in the alt text - for example 'Voluntary Arts logo'.

If images are purely for decorative purposes (borders, page backgrounds etc) then they should ideally be defined as 'background' in the PDF structure. This option is currently unavailable in OpenOffice Writer so an alternative is to give decorative images the alt text of 'background' until the issue is resolved in the software.

Hyperlinks

Using clickable hyperlinks in your documents will help your readers to easily navigate within the document, or to external websites. To create the correct tag structure use the Insert > Hyperlink feature or allow OpenOffice Writer to create links automatically.

However, do not create a series of links to different linked destinations using the same text as the link (for example 'Click here for A', 'Click here for B', etc). Instead use link text which communicates the context. Some screen readers can be used in 'link mode', which means that the link text alone is read out. In the example this would result in a confusing series of 'click here'. A better structure in this instance would be 'Go to A', 'Go to B' with descriptions for A and B etc.

Tables of contents

If you are creating multi-page documents, consider creating a table of contents (ToC) to assist all readers in navigating the resulting PDF. As with most other features in OpenOffice Writer, make sure you use the software's inbuilt ToC creation and updating features in order to ensure that the resulting PDF is correctly tagged.

If you have used OpenOffice Writer's inbuilt heading structure, you can create a table of contents by choosing Insert > Indexes and Tables > Indexes and Tables, and then click the Index/Table tab.

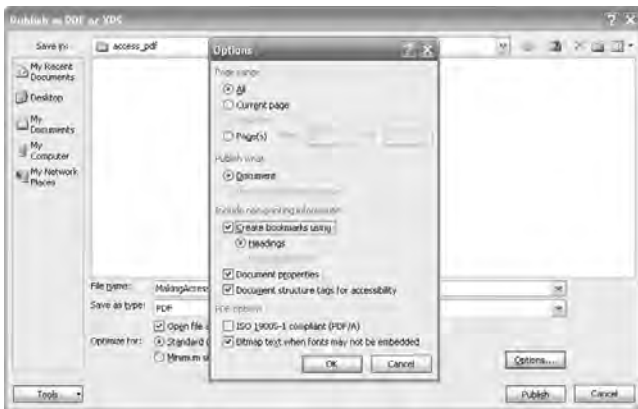


Save as PDF - stage 2.

To create accessible PDFs from your Word 2007 document:

- Click the Office button at top left, then choose 'Save as' > 'PDF or XPS'.
- In the next dialogue, choose your optimisation options (whether to reduce file size for online use) and then click 'Options'.
- Choose the options you require from 'Page range' and 'Publish what'.
- Under 'Include non-printing information' ensure that 'Document structure tags for accessibility' is checked. Also check 'Create bookmarks using' (choose 'Headings' if your document includes headings, or 'Word bookmarks' if you have created bookmarks manually).
- Click OK, then Publish to create your PDF.

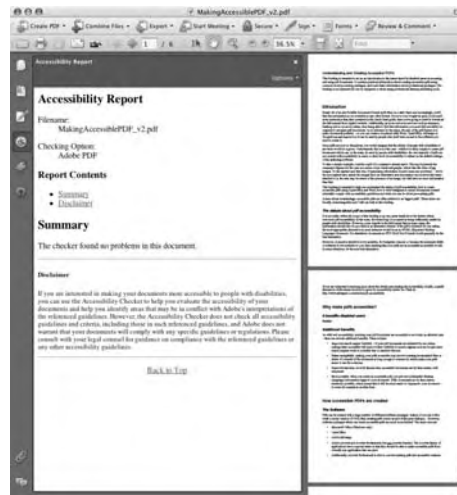
Ideally you should also check your PDF's accessibility using Acrobat Professional.



Setting accessibility options when exporting PDFs.

F. Commissioning accessible PDFs

If you use a designer to create your publications, make it clear when appointing them that any PDFs they produce need to be accessible. In most cases it is only the final PDF that needs to



An accessibility report produced by Acrobat Professional.

be tagged (rather than the PDFs the designer/company will send through whilst developing the design) but you should inform the designer of your wishes early on in the process as this may affect how they approach the work.

It should be sufficient to let the designer know that you want final PDFs to be 'accessible, tagged PDFs'. You might also consider requesting that each PDF be checked for accessibility using Adobe Acrobat Professional, and any errors corrected. Acrobat Professional will produce an accessibility report when it checks PDF accessibility, so these reports could become part of the design workflow if required.

Creating well structured, accessible PDFs does take more time than creating standard PDF documents, so designers may need to charge more to do this work. However, if you are a regular client the process should become part of their workflow so costs should reduce over time.

G. Checking existing PDFs

At the moment, the only way to check if a PDF has been tagged correctly is to use Adobe Acrobat Professional. The accessibility checker built into Acrobat Professional will perform a number of tests on PDF documents, checking for the presence of tags, alternative text for images, a text language declaration (for example English, Welsh or French), that content is contained within the document structure (so that the document can be read by a screen reader in a logical order) and that tables and forms are correctly marked up.

The software will produce a report outlining any problems in the document, or that it found no problems. However, as in web accessibility, automated checking is not the end of the story - the automated checker can only check for the

presence of alternative text and logical document structure, not whether that alternative text is accurate or the document structure correct.

H. Making existing PDFs accessible

If your organisation uses PDFs extensively, then it is likely that you will have a backlog of inaccessible PDFs which need to be made accessible. There are two ways to go about doing this:

1. Go back to the file and software which created the PDF originally and do the remedial work to create an accessible PDF. This may be the best solution if the file is regularly updated (for example a newsletter produced in Word).
2. Use Adobe Acrobat Professional to add a tag structure to the PDF. Ordinarily you will then need to do a considerable amount of manual adjustment of the tags, as well as manually adding alt text to images.

If the PDF you are trying to add tags to is not comprised of machine readable text (for example, if text is presented in the form of images) you may need to run the PDF through an Optical Character Recognition (OCR) program first, or add the alt text manually.

Acrobat Professional is currently the only software which can add tags to existing, untagged PDF documents.

Further Resources

Acknowledgements

VAN would like to acknowledge and thank artist and designer **Stefhan Caddick** for his expertise and help in creating this briefing. Further information on Stefhan's work is available at www.papergecko.co.uk.

Creating PDF forms

This briefing does not discuss the use and accessibility of PDF forms. To create accessible PDF forms you would need to use Adobe Acrobat Professional. Adobe has provided detailed information about this online at:

<http://acrobatusers.com/tutorials/creating-accessible-pdf-forms-acrobat-9-pro>.

The debate about PDF accessibility

A useful, detailed discussion on the debate surrounding PDF accessibility is given by accessibility writer Joe Clark at www.alistapart.com/articles/pdf_accessibility.

Tools

- Adobe Acrobat - www.adobe.com/products/acrobat
- OpenOffice - www.openoffice.org
- OpenOffice Writer - www.openoffice.org/product/writer.html

Information contained here may go out of date and you are therefore advised to check its currency. Updated information may be available on the VAN website: www.voluntaryarts.org

Disclaimer: Reasonable precautions have been taken to ensure the information in this document is accurate. However, it is not intended to be legally comprehensive; it is designed to provide guidance in good faith at the stated date but without accepting liability.

We therefore recommend you take appropriate professional advice before taking action on any of the matters covered herein.

© **Copyright notice:** Unless otherwise stated, all material published by VAN is subject to copyright.

However, we do encourage members of the voluntary arts sector to copy and disseminate this material for non-commercial purposes.

Prior to doing so, please send details of your requirements to info@voluntaryarts.org.

We also require that you acknowledge VAN wherever such material is used.

**Do you need this publication in an alternative format?
Contact T: 029 20 395 395 E: info@voluntaryarts.org**



The Voluntary Arts Network, Ground floor, 121 Cathedral Road, Pontcanna, Cardiff CF11 9PH

T: 029 20 395 395 **E:** info@voluntaryarts.org **W:** www.voluntaryarts.org

The Voluntary Arts Network is registered in Scotland as Company No. 139147 and Charity No. SC 020345.

VAN acknowledges funding from the Arts Councils of England, Northern Ireland, Scotland and Wales, and the Big Lottery Fund.

Registered office: 2nd Floor, 54 Manor Place, Edinburgh EH3 7EH.