



100% web based publishing platform that converts PDF files into eBooks and HTML. The digital files can be embedded into an existing website via an Iframe or linked back directly to the file.

Hosted and non hosted options are available.

Useful for district ten year plans, annual plans and other local government documentation. Can also be used for annual reports, product brochures etc

PDF Creation and Compatibility

The PDF conversion tool is designed to work with Adobe PDF files. However there are many formats and versions of these and whilst the majority of documents will convert to eBook and HTML, there are some compatibility issues with HTML which may cause some alignment and formatting issues.

When a PDF is created, text is normally kept as text, some scanned information is turned into images and some into text. Most of the text will be compatible with HTML text and is converted accordingly, however there are situations where this can't happen. For example, if the PDF converter finds text over an image it will keep it as text, even if the text is on an angle. HTML does not handle text on an angle so the HTML version will not look the same as the original PDF.

Frequently Asked Questions

How is scanned text rendered

This is determined by how Adobe does it in the PDF but usually in OCR or graphics or a mixture of both.

Can the tool convert from Word format (.doc, docx)?

This is not available in the current version of Digidocs. We are looking at offering this in a future version for organizations that require it. Whilst there are several Word to PDF converters only a few contain the functionality and interface that we require for automated conversion.

What is the target weight of an HTML page after conversion?

This depends entirely on the individual document and type of content – i.e. text or images. The pages themselves are small but if the document has very large images these will still be the same size in HTML.

How are the Flash pages (eBook) rendered?

They are rendered as SWF files.

What are some typical sizes for the flash versions

The conversion process creates two SWF files per PDF page (one for normal viewing, one for zoomed viewing). The size of these files varies from approximately 20Kb to 120Kb depending on page content.

Has there been any accuracy testing?

Yes there has been a large amount of testing carried out by Cucumber. From our testing we have found that the tool should convert most documents (excluding PDF version 1.5) but the HTML output and accuracy of display may vary depending on the content, PDF version and shortcomings between the PDF and HTML formats. We are working hard on trying to improve this for future versions. If you do experience any problems with individuals documents please contact support@cucumbersoftware so we can investigate.

How long does it take to convert a document

There are many variants that impact on the conversion time including type of document and content within the document as well as the hardware. Times from 10 seconds to 15 minutes should be expected, dependant on the document size.

Can the tool render HTML as text from a scanned document?

Yes, if the PDF has used OCR conversion and created it as text. Typically printed documents will be ok, but handwritten pages may be created as images. To check this, when you view a PDF the text is normally selectable but images aren't.

Digidocs handles the resultant PDF created from scanned documents. If the scan created an image it will be converted to an image in html, if the scan used OCR to keep text as selectable text then it will stay as text in the HTML.

Can the site be used to convert internal documents and external documents?

Yes it can, however it will need to be hosted outside in the DMZ .

Is there the ability to print the eBook version?

Yes, there is but we recommend using the PDF version.

Can we convert to only an eBook mode or an HTML mode not both?

Yes, this is configured at a site level.

When clicking on the link to open the document is it possible to link initially to the html version and then to the eBook version (as opposed to vice versa).

This is not currently set up but may be offered in the next version.

Why are the sections I've just created not visible in the drop down?

Your browser page needs to be refreshed as you are currently viewing an old version of the page stored in cache. Press Ctrl-F5 to refresh.

Editing HTML pages via the Digidocs content editing interface

While the HTML pages may be edited, care must be taken not to overwrite existing HTML that may not be visible in the content editor box except in source mode..

Current known issues with PDF Version 1.5

Currently Adobe PDF Version 1.5 is not supported by Digidocs as it will not convert to HTML. To check the version of a document, right click on name in document list and select the PDF tab.

Support

We are constantly trying to improve the tool, but if you do experience any difficulties converting documents please contact support -

support@cucumbersoftware.com

07 578 5022