

OUTRIDER WHITEPAPER

PDF OPTIMISATION

TABLE OF CONTENTS

TABLE OF CONTENTS	2
Scope	3
Objectives	3
What is PDF and why optimise PDF documents?	4
HTML or PDF	5
Tips for Creating Search Engine Friendly PDF documents.....	5
a. Use styles during creation of the PDF	5
b. Use the correct PDF writer.....	5
c. File naming conventions.....	5
d. Size of document.....	5
e. Embedded information	6
f. Include optimised content near the start of the document.....	6
g. PDF Reading Order	6
h. Add tags to your document	7
i. Version differences	7
j. Keep an eye on your changes	7
k. Calls to action	8
l. Sign ups	8
m. Include links to the document on the website.....	8
Resources	9
Confidentiality.....	10

Scope

The purpose of this document is to present a series of search engine optimisation recommendations for PDF files to maximize exposure on all major search engines and some tips for content design for PDF documents.

Most major search engines now serve PDF documents in the organic search results. Optimisation of PDF documents is as important as optimizing HTML pages. This whitepaper describes the optimisation techniques for PDF documents.

Objectives

Just as it is with webpage content, the fundamental objective of PDF documents should be to provide relevant information to the users and improve the brand awareness and recognition as the market leader.

Along with search engine optimisation, this document intends to provide a series of recommendations on creating user friendly PDF documents. These techniques are designed to ensure that the content is accessible and relevant to the users' search query.

What is PDF and why optimise PDF documents?

PDF stands for Portable Document Format. It was introduced by Adobe Acrobat with the intention of doing away with hardware and platform dependencies for documents which were designed to be printed on PostScript printers.

PDF has proved to be very useful over the years for sharing information in the form of newsletters, whitepapers, press releases, brochures, surveys and many others. Using PDF ensures the reliability of the format and look & feel of the document regardless of the computer used to access it.

As such the primary objective of PDF was to provide a uniform format through which documents could be shared. Online accessibility has never been a goal for PDF. However, search engines are evolving rapidly and have become better at indexing more and more content in various formats. Due to this optimising PDF content has become a mandatory requirement.

Google, for example, has about 455,000,000 PDF documents indexed in its search database. These documents find a way into the search engine results pages for search queries made by users based on the relevancy of the content.

Web Images Groups News more »

Google™ inurl:pdf Search Advanced Search Preferences

Search: the web pages from Australia

Web Results 1 - 10 of about 455,000,000 for inurl:pdf. (0.11 seconds)

[Page Request Not Found](#)
uscis.gov/graphics/formsfee/forms/files/t-9.pdf - 9k - [Cached](#) - [Similar pages](#)

[\[PDF\] 2006 Form W-4](#)
File Format: PDF/Adobe Acrobat - [View as HTML](#)
Form W-4 (2006). Purpose. Complete Form W-4 so that your employer can withhold the correct federal income tax from your pay. Because your tax situation ...
[www.irs.gov/pub/irs-pdf/fw4.pdf](#) - [Similar pages](#)

[\[PDF\] Form W-9 \(Rev. November 2005\)](#)
File Format: PDF/Adobe Acrobat - [View as HTML](#)
Give form to the requester. Do not send to the IRS. Form W-9. Request for Taxpayer Identification Number and Certification. (Rev. November 2005) ...
[www.irs.gov/pub/irs-pdf/fw9.pdf](#) - [Similar pages](#)

[PDF - Parkinson's Disease Foundation, Inc. | Home](#)
The **Parkinson's Disease Foundation** is a leading national presence in Parkinson's disease research, patient education, and public advocacy.
[www.pdf.org/](#) - 44k - [Cached](#) - [Similar pages](#)

[\[PDF\] The Not So Short Introduction to L T E X2 ε](#)
File Format: PDF/Adobe Acrobat
A. T. E. X2. ε. Or L. A. T. E. X2. ε. in 139 minutes. by Tobias Oetiker. Hubert Partl, Irene Hyna and Elisabeth Schlegl. Version 4.20, May 31, 2006 ...
[www.ctan.org/tex-archive/info/ishort/english/ishort.pdf](#) - [Similar pages](#)

[\[PDF\] The Digital Millennium Copyright Act of 1998](#)
File Format: PDF/Adobe Acrobat

Sponsored Links

[PDF Converter for MS Word](#)
New version just released! Quickly convert into MS Word. Download now.
[www.Nuance.com.au](#)

[PDF Online Printing](#)
Free Software to Create & Deliver Manage Entire Process from your PC
[PDF.BentleyPrinting.com.au](#)

[Top Quality EBooks Online](#)
Shop Sarahs CBmall 1900+ info books Easily find the Ebooks you want
[www.sarahscbmall.cjb.net](#)

[Fast Simple PDF Printing](#)
Free Software On-Line Now Deliver to 100's of Local Printers
[PDF.KKD.com.au](#)

[PDF Writer -Free Download](#)
Create PDF's from any application. Simple to Use, No Registration!
[www.docudesk.com](#)

[PDF Converter](#)

HTML or PDF

Although PDF documents are indexed by search engines, HTML pages with the same content are always preferable if ranking in the SERPs (Search Engine Results Pages) is a criterion for the success of the document. Where possible, create an HTML version of the content on the website. However, it is necessary to ensure that only the HTML version is available to the spiders to avoid duplicate content issues. Using the robots.txt file is recommended to effectively manage this.

Tips for Creating Search Engine Friendly PDF documents

a. Use styles during creation of the PDF

Depending on the software used to create the original document, use the same optimisation techniques as you would for an HTML page. Use styles for header tags, keywords in headings and sub-headings, and keywords in links. Pay attention to the keyword density and use image optimisation techniques just as you would for HTML pages.

b. Use the correct PDF writer

There are many PDF writers available. While most PDF writers will convert the text in the original document as text, there are some PDF writers which will render all the content into images during the creation of the PDF document. This will make the content inaccessible to search engines. A quick test is to select the text in the created PDF document and copying & pasting it into notepad. If the text can be selected, copied and pasted then the text will be accessible to search engines.

c. File naming conventions

Similar to other online file and folder names, the PDF documents should be named using keywords and meaningful text. Use hyphens and keywords in the names as it is a user and search engine friendly way of naming files and folders.

d. Size of document

Be mindful of the size of the PDF document. If the document is too long it is probably not focused on one or two main topics. Ensure that each PDF document has an integral focus. For bigger topics, the PDF document should be broken into a

series of documents. Each series should be linked to the other using relevant keywords.

e. Embedded information

Using the provided tags in the PDF document properties will help to emphasise the relevance of the content. There are two main ways to do this: Document Properties and Document Metadata. Under the Description Menu there are several fields which are accessible by search engines which include: Title, Author, Description, Keywords, and Copyright. Relevant keywords should be used while setting these properties for each document. Some examples are shown in the table below:

Metadata	Bad	Good
Title	tbuySep06.pdf	Tips for buying used cars in Australia
Copyright	TCS	The Car Site © 2006
Description	PDF doc tips	A step by step guide for buying used cars in Australia

f. Include optimised content near the start of the document.

Search engine spiders read through the content of PDF documents and select information to display on search engine results pages. PDF documents offer less control of what is displayed in comparison to the meta description tag of HTML web pages yet you can still influence what appears. Try to place several optimised sentences near the start of the PDF document as it is likely that these sentences will appear as the description within the search engine result pages.

g. PDF Reading Order

Every PDF document includes what is known as 'Reading Order'. The reading order of a PDF document indicates to users and search engines alike which areas of text are the most important.

You can enable the visual reading order tags by opening a PDF with Acrobat Writer and selecting the following options:

- Select Advanced
- Then Accessibility
- Then select Add Tags to Document

You will need to then select the following options to modify the reading order of a PDF document:

- Select Advanced
- Then Accessibility
- And finally, Touch Up Reading Order

The reading order of the PDF will be displayed and you will be able to indicate higher priority text areas within the document. You should place higher priority to content areas that contain keywords you are targeting.

h. Add tags to your document

Similar to HTML tags, PDF documents allow the placement of tags on such areas as headings and images. Acrobat produces a document report and will recommend areas that you may want to consider modifying. You can access this by opening the PDF document in Acrobat Writer and following these steps:

- Select Advanced
- Then choose Accessibility
- And finally select Add Tags to Document

This will produce a report and will give you the option of tagging headings, alternate text for images and more within your PDF document.

i. Version differences

While search engines do index PDF documents, it is wise to export your PDF documents as a version that is used by the majority of users. Whilst the latest version of Acrobat offers new features, often these are not supported by search engines. Users may not have the latest version and it is good practice to use the previous version to the latest available. Saving your PDFs using the last-but-one version of Acrobat should ensure search engines and users are able to access the content within your PDF documents easily.

j. Keep an eye on your changes

Each time you open a PDF document and make a small change then save it, modifications can occur. Often times the reading order can change automatically or you may inadvertently save the document as a higher version. It may save using the default settings rather than an optimised file size. It is wise to ensure each time you save a PDF that your optimisation settings are saved correctly.

k. Calls to action

Adding relevant links from the PDF content and footer to the website will provide inbound traffic from the document. Be sure to include contact information and the URL of the website along with relevant keywords in the footer of each document. Any calls to action which are relevant to the content should be added to the document with a link to the appropriate page on the website. For example, in a document titled "Tips on purchasing a used car", add a link to the "Search for used cars" page in the footer.

l. Sign ups

One way to use PDF documents as a measurable call to action is to invite users to sign-up or submit their details before downloading the PDF document. In this way all downloads can be tracked and reported on. This method involves implementing some code to display the form when a request is made for the PDF file.

A simple and quick method is to request submission of an email address before downloading the document.

m. Include links to the document on the website

Add links to the PDF documents on the website using relevant keywords. If there is no HTML version of the same content then a link to the PDF documents should also be included in the HTML, Google and other sitemaps for usability and search engine accessibility.

Resources

Accessibility:

http://www.adobe.com/products/acrobat/access_booklet.html

<http://www.webaim.org/techniques/acrobat/>

http://www.alistapart.com/articles/pdf_accessibility

PDF search

<http://www.google.com/help/features.html#pdf>

<http://www.searchtools.com/info/pdf.html>

<http://searchengineland.com/070912-095906.php>

Confidentiality

This document is commercial in confidence and has been prepared by Outrider for the purposes of reviewing by our clients only. Viewing or copying of this document by any other party other than our clients and their immediate staff is restricted.

Any questions regarding this document should be directed to:

Outrider
15-19 Parraween Street,
Cremorne, NSW 2090, Australia

Phone: +61 2 8968 4000
Fax: +61 2 9904 5055
Email: consult@outrider.com
Web: <http://www.outrider.com.au>