pdftotext – Portable Document Format (PDF) to text converter (version 3.00) [options] converts Portable Document Format (PDF) files to plain text. Pdftotext reads the PDF file, and writes a text file, If is not specified, pdftotext converts to If is ´-', the text is sent to stdout. Pdftotext reads a configuration file at startup. It first tries to find the user's private config file, ˜/.xpdfrc. If that doesn't exist, it looks for a system-wide config file, /etc/xpdf/xpdfrc. See the man page for details. Many of the following options can be set with configuration file commands. These are listed in square brackets with the description of the corresponding command line option. Specifies the first page to convert. Specifies the last page to convert. Maintain (as best as possible) the original physical layout of the text. The default is to ´undo' physical layout (columns, hyphenation, etc.) and output the text in reading order. Keep the text in content stream order. This is a hack which often "undoes" column formatting, etc. Use of raw mode is no longer recommended. Generate a simple HTML file, including the meta information. This simply wraps the text in <pre> and </pre> and prepends the meta headers. Sets the encoding to use for text output. The must be defined with the unicodeMap command (see The encoding name is case-sensitive. This defaults to "Latin1" (which is a built-in encoding). Sets the end-of-line convention to use for text output. Don't insert page breaks (form feed characters) between pages. Specify the owner password for the PDF file. Providing this will bypass all security restrictions. Specify the user password for the PDF file. Don't print any messages or errors. Read in place of ˜/.xpdfrc or the system-wide config file. Print copyright and version information. Print usage information. and are equivalent.) Some PDF files contain fonts whose encodings have been mangled beyond recognition. There is no way (short of OCR) to extract text from these files. The Xpdf tools use the following exit codes: 0 No error. 1 Error opening a PDF file. 2 Error opening an output file. 3 Error related to PDF permissions. 99 Other error. The pdftotext software and documentation are copyright 1996-2004 Glyph & Cog, LLC.