

RTF Pocket Guide  
Keith Alcock, TCS and Developers SIG Member  
tcs@keithalcock.com

Microsoft's Rich Text Format, the file format used by Word primarily for document exchange, is the subject of this small volume, which is, as far as I know, the only book published on RTF. Readers may recall that Microsoft used it at one time for backward compatibility between Word versions in a rather underhanded way until it was discovered that the supposed \*.doc file was really a \*.rtf file in disguise. Microsoft was pressured to recant and in later versions of Word you will notice a file type called "Word 97-2002 & 6.0/95 - RTF (\*.doc)" that resulted.

Despite RTF's short-lived fame and more than a decade of behind the scenes responsibility, most people are only peripherally aware of it. The proclamation on the book's cover that RTF is the "Universal Document Format" is a difficult one to justify presently. It has long been surpassed by HTML/XML and maybe even PDF, in my opinion. Word will not entrust everything to a \*.rtf file that it will store in a \*.doc file and some versions of Word do not even adhere to Microsoft's own specification. So, don't fret that you missed the technology boat and buy this book to update your skills and land that lucrative job. If you are a programmer, however, and have to generate or process RTF, then this book is a valuable introduction to the topic that can smooth the steep learning curve associated with the definitive document on the subject, Microsoft's 248-page RTF specification, and it is well worth reading.

The book is divided into four parts of several chapters apiece: RTF Tutorial, Creating MS Windows Help Files, Example Programs, and Reference Tables. Part one occupies more than half the book and is unquestionably the most beneficial. It explains RTF syntax, dividing it into four categories: commands, escapes, groups, and text. Be prepared to deal with regular expressions here and in the example programs. The author covers the most useful commands for formatting paragraphs, characters, columns, and tables and well as drawing lines, embedding images, specifying colors, and several others. Applying information from this part and without reference to the RTF specification, I was able to write a program to extract commands from an RTF file and another to reformat the file to show its command structure in human readable format. This alone made reading the book worthwhile.

The book's website includes an explanation of how the book's structure evolved. The remaining three parts were originally appendices and perhaps should have remained so. The Windows help files it describes are the WinHelp variety vintage 1995 or before rather than the more recent HTML Help or emerging Windows Help 2 files, although URLs to more up-to-date (and expensive) products are provided. Older material is atypical of O'Reilly books which generally cover cutting edge topics. The example programs are all written in perl, which makes RTF downright easy to read in comparison. They include a datebook generator, directory lister, origami CD case maker, and RTF metadata extractor. The code can be downloaded for free, so it

does not warrant purchasing the book. Reference tables include the RTF character chart, RTF language codes, and a twips to point, inch, and centimeter converter.

A more comprehensive volume might require the unenviable title "RTF Manual." Nevertheless, additional topics would round out the book. Documentation on how to use Microsoft's Rich Textbox ActiveX component to generate RTF is one of them. Another is the difference between the RTF that Word and WordPad understand. Word can be turned into an RTF report generator via VBA code or calls to Word's COM interface, which is an interesting possibility. Tabs are not covered at all in the text even though they, along with space characters, help set RTF apart from HTML. Also, RTF commands like \ftnbj and \chshdng0 are difficult to understand and remember without a translation table.

If you really need to know about RTF and dread reading the Microsoft specification, then this book is a recommendable introduction. If you have not previously heard of RTF, rest assured that it will probably stay that way.

Title: RTF Pocket Guide  
Author: Sean M. Burke  
Publisher: O'Reilly  
Pages: 149  
Price: US \$12.95  
ISBN:0-596-00475-3  
Website: <http://www.oreilly.com/catalog/rtfpg/>

Keith Alcock is a software developer and consultant interested in many topics including computational linguistics, automated text reformatting, and spell checking.