

RTF2HTML V2.1

FUNKTIONALE BESCHREIBUNG.....

NICHT ÜBERSETZTE RTF-FEATURES.....

MÖGLICHE FEHLERQUELLEN.....

PROGRAMM-ENTWICKLUNGS-VERLAUF.....

Funktionale Beschreibung

RTF2HTML übersetzt RTF-Files aller RTF-Versionen nach HTML 3.

Aufruf-Parameter:

- 'optimize'
eliminiert überflüssige HTML-Tags wie z.B. '' oder '</SUB><SUB>'
- 'onlyDefiniteOpt'
sorgt dafür, das Strings wie '' NICHT wegoptimiert werden, da das schließende ''-Tag u.u. eine andere Anweisung als '' hier im Beispiel deaktivieren könnte
- 'noFonts'
deaktiviert alle '-Anweisungen

Folgende Formatierungen werden, so weit im HTML 3 möglich, äquivalent übersetzt:

- RTF-Stylesheets: die darin gespeicherten Formatierungen werden in die Übersetzung mit einbezogen
- Zeichenformatierungen: fett, kursiv, unterstrichen, durchgestrichen, tiefgestellt, hochgestellt
- Absatzformatierungen: zentriert, links- und rechtsbündig
- Aufzählungen aller Arten
- left indents (mittels -Schachtelungen)
- Zeilenumbruch / Seitenumbruch (als <HR>) /Absatz
- etwaige Farb-/Schriftart-/Schriftgröße-Formatierungen
- Sonderzeichen (' " < > usw.)
- Tabellen

Nicht übersetzte RTF-Features

Folgendes wird in der vorliegenden Version ignoriert:

- Kopf-/Fußzeile
- File Tables
- Bilder (im HTML-Dokument substituiert durch den String „[*** picture ***]“)
- bestimmte rtf-spezifische Formatierungen (z.B. kerning, right indent u. dgl.)
- Dokument-Infos (Autor, letzte Speicherung etc.)

Mögliche Fehlerquellen

Folgendes kann Fehler bzw. unerwünschte Ergebnisse verursachen:

- Der Aufrufparameter 'optimize' bewirkt, daß auch Zeichenketten wie '' gnadenlos wegoptimiert werden kann, was leicht in Formatierungs-Fehlern (NICHT HTML-Syntax-Fehlern) enden kann; Abhilfe: Parameter 'onlyDefiniteOpt'
- RTF-Dokumente sind oft wenig konsistent gespeichert; wenn z.B. Aufzählungen & Einrückungen einmal mit einem Einzug, dann wieder mit Tabulator X und dann vielleicht mit einem Aufzählungs-Befehl gespeichert werden, wird das HTML-File aller Wahrscheinlichkeit nach kleinere Fehler bei den Längen der Einrückungen und Tabulatoren aufweisen
- Übernahme von Text-Formatierungen in eine Tabelle, wenn eine solche beginnt, was in RTF rein theoretisch möglich ist, findet NICHT statt
GRUND: 1. werden beim Beginn einer Tabelle normalerweise ohnehin alle Text-Formatierungen zurückgesetzt
2. müßte man eine 'mitgeschleifte' Formatierung in einer HTML-Tabelle Feld für Feld neu setzen
und am Feld-Ende wieder löschen → das ausgespuckte HTML-File wird ****SEHR**** groß
- Aufzählungen in Tabellenfeldern (eher selten) werden nur mit Pseudo-Tabs und ·'s übersetzt (ohne ,)
- wenig sinnvolle RTF-Dokumente mit Punkten im Inhaltsverzeichnis, zu denen keine entsprechende Überschrift existiert verursachen HTML-Dokumente mit Phantasie-Referenzen

Programm-Entwicklungs-Verlauf

V 1.0: - erste offizielle Version

V 1.1: - Änderung bei der Behandlung von Gruppen-Enden
- Änderung bei der Darstellung von Bullet-Listen

V 1.2: - Übersetzung von Tabellen
- erweiterte Sonderzeichen-Behandlung
- geänderte Behandlung von Aufzählungen
- kleine Layout-Bereinigungen

V 1.3: - Änderung bei der Behandlung von Gruppen-Anfängen
- neue Prozedur zur Steuerung der Ausgabe von Dokument-Text
- selten auftretender Zeichenformat-Bug entfernt

V 2.0: - Einbindung von Stylesheets
- Inhaltsverzeichnis/Überschrift-Verweis-Strukturen werden in HTML-Sprungmarken umgewandelt
- Wörter, die mit 'http://' beginnen, werden automatisch in Hyperlinks umgewandelt
- Aufzählungen (auch geschachtelt) werden als entsprechend strukturierte 's nach HTML konvertiert

- verbesserte HTML-Code-Optimierfunktion
- Aufrufparameter für rtf2html() zum Variieren der Konvertier-Vorgangsweisen
- diverse kleine Layout-Änderungen

V 2.1: - überarbeiteter Formatierungs-Algorithmus
- Übersetzung von left indents