

Word-Tipp: HTML-Tags löschen

Versionen: Word 97 bis 2003

Wenn Text aus HTML-Seiten in bereits vorhandene Dokumente übertragen oder eine HTML-Seite als Basis für ein neu zu erstellendes Dokument verwendet werden soll, müssen die in HTML-Seiten enthaltenen «Tags» wie «<head>», «p>» oder «table>» mühsam manuell gelöscht werden, wobei nicht selten irrtümlich Text gelöscht wird, weil er zwischen den HTML-Anweisungen nur schwer zu erkennen ist.

Sinnvoll ist es deshalb beim Entfernen der HTML-Tags die Suchen/Ersetzen-Funktion von Word zu nutzen.

1. Wählen Sie im Dokument mit den HTML-Tags das Menü Bearbeiten-Ersetzen an.
2. Geben Sie im Feld Suchen nach die folgende Zeichenfolge ein:
`\<*\>`

Das Feld Ersetzen durch bleibt leer!

3. Aktivieren Sie das Kontrollkästchen Mit Mustervergleich (Word 97 und 2000) oder Platzhalter verwenden (Word 2002/2003). Wenn dieses Kontrollkästchen im Ersetzen-Dialog nicht vorhanden ist, klicken Sie auf die Schaltfläche Erweitern. Word blendet dann zusätzliche Einstellungen unter den Eingabefeldern ein.
4. Klicken Sie auf die Schaltfläche Alle ersetzen.
5. Bestätigen Sie die abschliessende Erfolgsmeldung von Word mit Ok.

Die Zeichenfolge «\<*\>» bedeutet im Klartext: Suche nach Textpassagen, die mit dem Sonderzeichen «\<» (Spitze Klammer auf) beginnen, bei denen es keine Rolle spielt, was für Zeichen dann folgen (Platzhalter «*»), die aber mit dem Sonderzeichen «\>» (Spitze Klammer zu) enden sollen. Das Sternchen «*» steht hier als Platzhalter für beliebige Zeichen in den spitzen Klammern, wodurch alle HTML-Tags erfasst sind. Gelöscht werden die HTML-Tags, weil das Feld Ersetzen durch leer gelassen wurde.

In einzelnen Zeilen der HTML-Seiten sind die Tags der einzige Inhalt. Nachdem Sie die Tags wie oben erläutert gelöscht haben, bleiben also viele leere Zeilen, die ebenfalls entfernt werden müssen. Dazu benutzen Sie wieder die

Ersetzen-Funktion: Im Feld Suchen nach geben Sie dazu «^a^a» (Word 97) bzw. «p^p» (Word 2000, 2002/XP und 2003) und im Feld Ersetzen durch geben Sie «^a» (Word 97) bzw. «^p» (Word 2000, 2002 und 2003) ein. Deaktivieren Sie das Kontrollkästchen «Mit Mustervergleich» (Word 97 und 2000) oder «Platzhalter verwenden» (Word 2002/XP und 2003). Nach einem Klick auf «Alle ersetzen» tauscht Word zwei aufeinander folgende Leerzeilen gegen eine Leerzeile aus. Wiederholen Sie den Ersetzen-Vorgang mit diesen Einstellungen, bis Word nur noch eine einzelne Ersetzung meldet. Damit sind alle überflüssigen Leerzeilen gelöscht und Sie können mit der weiteren Bearbeitung der HTML-Seite fortfahren.

www.wdsch.de