

Word-Dokument wird webfit

Word-Dokumente sind weit verbreitet: als Skripten, als Buchvorlagen als Unterrichtsbehelfe. Wenn sie auch im Internet publiziert werden sollen, gibt es die übliche Methode der in Word eingebauten Umformungsfunktion (Speichern unter...), die eine Html-Version erzeugt. Speziell bei langen Dokumenten verursacht diese Html-Version aber große Ladezeiten und bietet keine Navigationshilfen. Dieser Beitrag zeigt, wie man mit Bord-Mitteln von Microsoft Word 2003 rasch ein navigierbares Webdokument herstellen kann.

Franz Fiala

Word erlaubt die Speicherung von Originaldokumenten sowohl als Doc-Datei als auch als Html-Datei mit der Funktion **Datei** ▶ **Speichern unter**... Insofern bereitet das Publizieren von Word-Dokumenten im Html-Format keine Probleme.

Allerdings werden an Html-Dokumente andere Anforderungen gestellt als an Dokumente, die ausgedruckt werden. Insbesondere benötigt man speziell bei großen Dokumenten Navigationselemente, die ein Blättern zwischen den einzelnen Abschnitten des Dokuments erleichtern.

Dabei bietet Word einige Werkzeuge an, die auch für das Publizieren im Html-Format vorteilhaft verwendet werden können. Neben dem einfachen Speichern als Html-Datei unterstützt Word auch

- Frames (**Format** ▶ **Frames**),
- die Aufteilung eines großen Dokuments in mehrere kleine Filialdokumente (Unterdokumente) (**Gliederung** ▶ **Unterdokument erstellen**)

und erlaubt in jeder dieser Versionen immer auch das Speichern im Html-Format. Daher ist Word für die Publikation einer Doc-Datei als Web-Dokument ganz gut gerüstet.

Wenn man aber die weitergehenden Features wie Frames oder Filialdokumente nutzt, stellt man fest, dass die entstehenden Html-Versionen nicht immer ohne händische Nacharbeit verwendbar sind. Word ist eben nicht in erster Linie für Html-Dokumente konzipiert.

Bei jeder Änderung des Originaldokuments muss auch die Html-Version geändert werden. Handelte es sich dabei nur um einen einfachen Speichervorgang (**Speichern unter**...), wäre das weiter kein Problem. Wenn aber zur Erzeugung der Html-Seite arbeitsintensive Nachbearbeitung notwendig wird, dann ist eine zügige Weiterentwicklung behindert und die Folge ist dann, dass bei Änderungen am originalen Word-Dokuments die Umwandlung in das Html-Format unterbleibt und sich Doc- und Html-Version auseinander entwickeln.

Hier wird beschrieben, wie man die vorhandenen Elemente von Word zur Herstellung leichter lesbarer Html-Dokumente ausnutzen kann und mit möglichst wenig Handarbeit auch eine gut navigierbare Html-Version erzeugt.

Solche Word-Projekte sollten immer in einem eigenen Verzeichnis gespeichert werden, weil bei der Verwendung von Frames, Filialdokumenten und dann auch noch der Html-Version nicht mehr nur ein sondern viele Teildokumente existieren.

Einstellungen

Bevor ein Word-Dokument in eine Html-Version konvertiert wird, sollte man die zugehörigen Optionen (**Extras** ▶ **Optionen** ▶ **Allgemeines** ▶ **Weboptionen**) studieren und die einzelnen Kategorien **Browser**, **Dateien**, **Bilder**, **Codierung** und **Schriftarten** einstellen.

Der "Wortschatz" wird unter **Browser** eingestellt. Voreingestellt ist **Microsoft Internet Explorer 4.0 oder höher**, man kann aber auch andere Versionen wählen und Navigator-Kompatibilität einstellen.

Unter **Dateien** wird festgelegt, dass Hilfsdateien in einem eigenen Ordner gespeichert werden sollen. Das ist übersichtlich und sollte jedenfalls gewählt werden.

Unter **Bilder** wird die Auflösung der gespeicherten Bilder eingestellt, was sowohl die Ladezeit aber auch die Bildqualität bei der Darstellung beeinflusst.

Unter **Schriftarten** können der Zeichensatz sowie die Proportional- und die Festbreitenschrift und deren Größe eingestellt werden.

Datei ▶ Speichern unter...

Um ein Word-Dokument in eine Html-Version zu konvertieren, verwendet man die Kommandofolge **Datei** ▶ **Speichern unter ...** mit den Speicheroptionen

- **Webseite** *.htm
- **Webseite gefiltert** *.htm
- **Webseite in einer Datei** *.mht
- **XML-Dokument** *.xml

Bei den beiden ersten Möglichkeiten wird bei Bedarf auch ein gleichnamiger Ordner mit dem Suffix **-Dateien** (oder **-Files** in der englischen Version von Word) generiert, in dem neben einer Dateiliste (**filelist.xml**) auch alle Bilder und gegebenenfalls auch Fußnotendateien enthalten sind.

Der Ordner ***-Dateien** entsteht nur bei Bedarf also etwa dann, wenn die Doc-Datei auch Bilder enthält. Sollten Sie diesen Ordner nicht sehen, müssen Sie im Windows-Explorer unter **Extras** ▶ **Ordneroptionen** ▶ **Ansicht** ▶ **Paare von Webseiten und Webordnern verwalten** umschalten auf **Beide Teile anzeigen, aber als einzige Datei verwalten**. Bei einer Publikation der Html-Version muss nicht nur die entstandene Html-Datei sondern auch dieser Ordner mit allen Inhalten publiziert werden.

Die Speicheroptionen sind:

Webseite

Das ist jene Version, die in der Html-Version das Layout der Doc-Datei möglichst genau nachbildet. In den meisten Fällen sind Doc-Version und Html-Version gleichwertig und

es genügt eine der beiden zu speichern und weiter zu entwickeln. Nachteil der entstehenden Dateien sind oft lange Ladezeiten und möglicherweise Ungenauigkeiten bei fremden Browsern.

Webseite gefiltert

Diese Html-Version verzichtet auf XML-Elemente und damit auf maximale Übereinstimmung von Original mit Html-Version, ist aber schneller und kommt einer üblichen Html-Version näher und wird in allen Beispieldateien dieses Beitrags verwendet.

Webseite in einer Datei

Hier werden die Daten aller beteiligten Dateien in einer einzigen Datei kodiert. Diese Version ist von Vorteil, wenn das Html-Konstrukt in einer Datenbank gespeichert werden soll.

Eigenschaften der Html-Datei

Eine so als Html gespeicherte Word-Datei kann in einem Browser nur über die vertikalen und horizontalen Scroll-Balken navigiert werden. Darüber hinaus haben große Word-Dokumente entsprechende Ladezeiten, und es empfiehlt sich, das Dokument in mehrere kleinere Dokumente aufzuspalten, die über ein Inhaltsverzeichnis navigiert werden.

Wenn man diese Aufteilung händisch durchführt, kann das ziemlich mühsam werden, weil jeder dieser Abschnitte einen Namen benötigt und daher muss das gemeinsame Inhaltsverzeichnis mit Hilfe der gewählten Namen händisch erstellt werden. Der größte Nachteil ist aber der, dass bei Änderungen am ursprünglichen Word-Dokument alle die-

Hinweise

Referenzen: Word bietet auch verschiedene Referenzfunktionen wie Inhaltsverzeichnis und Index an (**Einfügen** ▶ **Referenz**...). Originaldokumente mit solchen Referenzfunktionen können nicht als Filialdokumente gespeichert werden. Stattdessen wird bei Dokumenten, die für Internet-Publikation vorgesehen sind, mit der Frame-Funktion ein eigenes Inhaltsverzeichnis erstellt.

Beispieldateien: Die beschriebenen Dateien sind bei der Webversion dieses Artikel im Verzeichnis **_prg** sowohl als downloadbare Dateien als auch als Zip-Archiv gespeichert.

Verzeichnisse: Die in dieser Dokumentation verwendeten Verzeichnisse **_prg/0, 0a, 0b, 1, 1a, 2, 2a** werden bei einem realen Projekt nicht benötigt und dienen hier nur zur Dokumentation der einzelnen Arbeitsschritte.

se Gliederungsarbeiten zu wiederholen wären.

Originaldokument vorbereiten

Ausgangspunkt ist eine Datei _prg/0/Mein Text.doc (**Bild 1**); Kapitelüberschriften und Text sind zunächst gleichrangig (Format **Standard**).

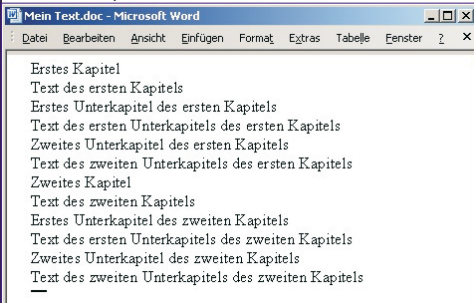


Bild 1: Rohtext im Format Standard

Das Dokument wird gegliedert. Den Überschriften der Kapitel werden Überschrifts-Formatvorlagen zugewiesen (**Bild 2**).

Text	Formatvorlage
Erstes Kapitel	Überschrift 1
Text des ersten Kapitels	Standard
Erstes Unterkapitel...	Überschrift 2
...	



Bild 2: Dokument in Normalansicht

Bei Bedarf wir mit **Format ▶ Nummerierung und Aufzählungszeichen ... ▶ Gliederung ▶ Überschrift** eine Nummerierung eingefügt.



Bild 3: Dokument mit Nummerierung

Man kann das Dokument sowohl in der Normalansicht als auch in der Gliederungsansicht (**Ansicht ▶ Gliederung**) (**Bild 4**) betrachten. In der Gliederungsansicht ist die Rangordnung der Abschnitte zu erkennen.

Frame

Ein Dokument wird mit Frames navigierbar:

- Kopieren des Dokuments _prg/0b/Mein Text.doc in das Verzeichnis _prg/1
- Einfügen eines Frame mit Inhaltsverzeichnis durch **Format ▶ Frames ▶ Inhaltsverzeichnis mit Frames**

Es entsteht ein Frame und ein Inhaltsverzeichnis im linken Frame, der Inhalt selbst bleibt im rechten Frame. Man kann jedes Ka-

pitel im linken Frame anklicken und danach scrollt der Text im rechten Fenster an die

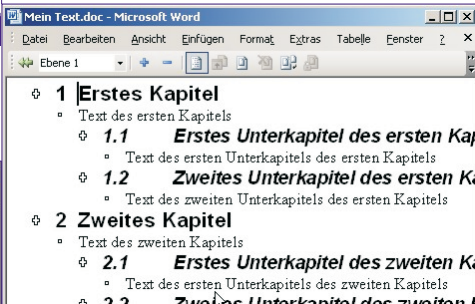


Bild 4: Dokument in Gliederungsansicht

richtige Stelle.



Bild 5: Dokument mit Frames

- **Datei -> Speichern -> default_Frame.doc**

Im Verzeichnis _prg/1a existieren jetzt drei Dateien:

Dokument	Aufgabe
default_Frame.doc	Frame-Dokument
Mein Text.doc	eigentlicher Inhalt
1Erstes Kapitel.doc	Inhaltsverzeichnis

Umbenennen einer Framedatei

Der Dateiname für das Inhaltsverzeichnis (hier 1Erstes Kapitel.doc) wird vom Text der ersten Zeile in der Gliederung bestimmt, sollte aber eigentlich Inhaltsverzeichnis.doc oder ähnlich heißen. So benennt man die Datei um:

- Kopieren aller Dateien aus dem Verzeichnis _prg/1 in das Verzeichnis _prg/1a
- Öffnen der Datei default_Frame.doc (**Bild 5**)
- Cursor im linken Frame
- Rechte Maustaste im oberen (oder unteren) Rand des Frame ▶ **Aktuellen Frame speichern unter ... ▶ Inhaltsverzeichnis.doc**
- Rechte Maustaste im oberen (oder unteren) Rand des Frame ▶ **Frameeigenschaften ▶ Anfangsseite ▶ Inhaltsverzeichnis.doc**
- **Datei ▶ Speichern**
- Datei 1Erstes Kapitel.doc löschen

alt	neu
1Erstes Kapitel.doc	Inhaltsverzeichnis.doc

Erzeugen der Html-Dateien

Durch **Datei ▶ Speichern unter... ▶ Webseite gefiltert** erhalten wir im Verzeichnis _prg/1a die Dateien

Dokument	Aufgabe
default_Frame.htm	Frame-Dokument
Mein Text.htm	eigentlicher Inhalt
Inhaltsverzeichnis.htm	Inhaltsverzeichnis

Ein Klick auf einen Link im Inhaltsverzeichnis navigiert den Inhalt auf die Position des betreffenden Kapitels. Der Inhalt ist in der einzigen Datei Mein Text.htm enthalten, was eventuell lange Ladezeiten zur Folge hat.

Frame mit Detaildokumenten

Ein Word-Dokument kann in ein Masterdokument und beliebig viele Filialdokumente aufgeteilt werden. Das Masterdokument hat eine ähnliche Funktion wie ein Frame-Dokument, kann aber darüber hinaus auch Texte enthalten. Die Filialdokumente können unabhängig vom Masterdokument bearbeitet werden.

Vorgangsweise

- Kopieren aller Doc-Dateien aus dem Verzeichnis _prg/1a in das Verzeichnis _prg/2
- Öffnen der Datei Mein Text.doc
- **Ansicht ▶ Gliederung**
- Überschrift und den nachfolgenden Text, der in einem getrennten Dokument gespeichert werden soll, markieren
- erzeugen eines Filialdokuments mit einem Klick auf das Symbol **Unterdokument erstellen** in der Symbolleiste **Navigation**
- Wiederholung für alle Kapitel

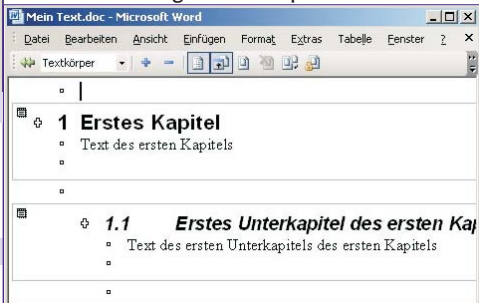


Bild 6: Erzeugen von Filialdokumenten

- Durch **Datei ▶ Speichern** werden die Filialdokumente (Detaildokumente) angelegt. Für das Beispieldokument Mein Text.doc erhält man folgende Dateien:

Dokument	Aufgabe
default_Frame.doc	Framedokument
Inhaltsverzeichnis.doc	Linker Frame
Mein Text.doc	Masterdokument
Erstes Kapitel.doc	Filialdokument
Erstes Unterkapitel im ersten Kapitel.doc	Filialdokument
Zweites Unterkapitel im ersten Kapitel.doc	Filialdokument
Zweites Kapitel.doc	Filialdokument
Erstes Unterkapitel im zweiten Kapitel.doc	Filialdokument
Zweites Unterkapitel im zweiten Kapitel.doc	Filialdokument

Die Namen der Filialdokumente wurden automatisch von den Kapitelüberschriften abgeleitet.

Wenn man jetzt das Masterdokument Mein Text.doc schließt und erneut lädt, bekommt man eine geänderte Ansicht, die den Pfad der Filialdokumente anzeigt:

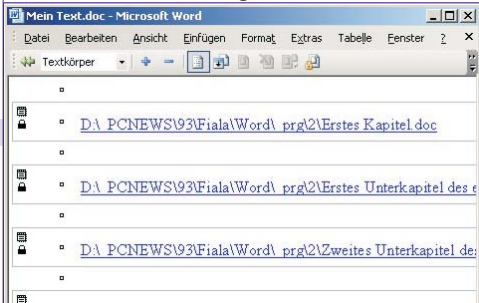


Bild 7: Pfade der Filialdokumente

http://demo.pcc.cc/ -> WebWord

Mit dem Symbol **Unterdokumente erweitern** in der Symbolleiste **Gliederung** kann man die Texte der Filialdokumente wieder wie in **Bild 6** sichtbar und bearbeitbar machen.

Mit der Taste **[Alt] [F9]** kann man den Code hinter der Darstellung sichtbar machen; das Schloss-Symbol deutet darauf hin, dass die Pfadangaben nicht verändert werden können.

Dateinamen

Die vorgegebenen Namen der Filialdokumente sollte man ändern, weil die langen Bezeichnungen in einem Html-Link sehr unübersichtlich sind.

- Kopieren aller Dateien aus dem Verzeichnis `_prg/2` in `_prg/2a`
- Öffnen der Datei `_prg/2a/Mein Text.doc`
- Klick auf den Pfadnamen bei gedrückter **[Strg]**-Taste öffnet das Filialdokument
- Speichern des Filialdokuments unter dem neuen Namen
- Wiederholen für alle Filialdokumente
- Speichern des Dokumentes `Mein Text.doc`
- Löschen aller alten Filialdokumente

alt (wird gelöscht)	neu
Erstes Kapitel.doc	A.doc
Erstes Unterkapitel im ersten Kapitel.doc	Aa.doc
Zweites Unterkapitel im ersten Kapitel.doc	Ab.doc
Zweites Kapitel.doc	B.doc
Erstes Unterkapitel im zweiten Kapitel.doc	Ba.doc
Zweites Unterkapitel im zweiten Kapitel.doc	Bb.doc

Öffnet man die Datei `default_Frame.doc` mit Word erhält man folgendes Bild:



Bild 8: `default_Frame.doc`

Erzeugen der Html-Dateien

Die Internet-Version erhält man durch:

- **Datei ▶ Speichern unter... ▶ Webseite gefiltert.**
- Die Frage, ob das Original überschrieben werden soll, beantwortet man mit **Ja, alle.**

Öffnet man die Datei `default_Frame.htm` mit Word erhält man folgendes Bild:

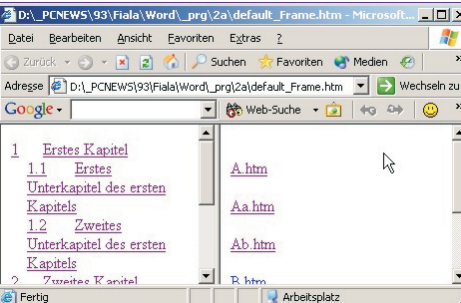


Bild 9: `default_Frame.htm`

Leider ist das Resultat nicht genau das, was man gerne hätte.

• Die Datei `Mein Text.htm` zeigt als Text nicht den Text des Filialdokuments (*Erstes Kapitel*) sondern das Masterdokument mit den Namen der Filialdokumente (*A.htm...*).

• Die Links im Inhaltsverzeichnis verweisen auf das Masterdokument und nicht auf die Filialdokumente.

Jetzt muss nach einer Möglichkeit gesucht werden, mit geringstmöglichem Aufwand an Handarbeit eine Situation herstellen, die es auch später mit vertretbarem Aufwand ermöglicht, aus dem Word-Original im Verzeichnis `_prg/2a` wieder eine navigierbare Html-Version herzustellen.

Die beiden Dateien, die die genannten Mängel enthalten werden unter neuen Namen kopiert und editiert. Wenn daher später die Html-Version noch einmal hergestellt wird, werden zwar die ursprünglichen Dokumente überschrieben, die kopierten und bearbeiteten Versionen bleiben aber unverändert und man muss das Editieren nicht noch einmal wiederholen, schlimmstenfalls geringfügig ändern.

Dokument	kopieren auf
<code>default_Frame.htm</code>	<code>default.htm</code>
<code>Inhaltsverzeichnis.htm</code>	<code>Inhalt.htm</code>

Jetzt werden die neuen Dateien mit einem Texteditor editiert:

Änderungen in `default.htm`

• `src`-Attribut des Inhaltsverzeichnisframe (Frame 2) bekommt den Wert `Inhalt.htm` (und zeigt damit auf das neue editierte Inhaltsverzeichnis)

• `src`-Attribut des Inhaltsframe (Frame 1) bekommt den Wert `A.htm` (und zeigt damit beim Öffnen auf das erste Kapitel)

Änderungen in `Inhalt.htm`

Die `href`-Attribute werden sinngemäß geändert:

- `alt`: `1 Erstes Kapitel`
- `neu`: `1 Erstes Kapitel`

Arbeitstechnik bei Änderungen

• Löschen aller Html-Filialdokumente im Verzeichnis `_prg/2a` nicht aber der Dateien `Inhalt.htm` und `default.htm`. (Das Löschen der Dateien ist erforderlich, weil bei jedem Überschreiben der vorigen Version die neueren Dateien andere Dateinamen mit Indexnummern erhalten würden. Damit wären die händisch editierten Links falsch.)

• Die Originaldateien sind die Doc-Dateien im Verzeichnis `_prg/2a`.

• Durchführung der Änderungen an den Doc-Dateien im Verzeichnis `_prg/2a` (ohne Änderung der Gliederung)

- Filialdokumentansicht einstellen
- Inhaltsframe aktivieren

• Speicherung als Html-Datei `default_Frame.htm` im Verzeichnis `_prg/2a`. Dabei werden alle Filialdokumente neu angelegt. Die bestehenden Dateien `Inhalt.htm` und `default.htm` werden bei diesem Vorgang nicht verändert.

Kurzfassung

Originaldokument vorbereiten

- Das Dokument sollte vor diesen Arbeiten in einer einigermaßen endgültigen Version vorliegen (`_prg/0/Mein Text.doc`)
- Man gliedert das Dokument in Kapitel und Unterkapitel (`_prg/0a/Mein Text.doc`) und fügt gegebenenfalls eine Überschriften-Nummerierung ein (`_prg/0b/Mein Text.doc`)

Frame

In vielen Fällen genügt es, dem Dokument, zur Navigation ein webfähiges Inhaltsverzeichnis hinzuzufügen.

- Erweitern des Dokuments um ein Inhaltsverzeichnisframe (`_prg/1/*.doc`)
- Umbenennen der entstehenden Inhaltsverzeichnisdatei (`_prg/1a/*.doc`)
- Herstellen der Html-Version mit **Speichern unter...** (`_prg/1a/*.htm`).

Nachbearbeitung

Als Originaldatei dienen die Doc-Dateien im Verzeichnis `_prg/1a/*.doc`. Diese Dateien werden bei Bedarf modifiziert. Bei Änderungen im Inhaltsverzeichnis muss auch das Inhaltsverzeichnis neu aufgebaut werden.

Zum Herstellen der Html-Version genügt der Befehl **Speichern unter...**

Frame mit Filialdokumenten

Bei großen Dokumenten ist es notwendig, das Dokument in mehrere Teildokumente aufzuteilen.

- Umwandlung der Kapitel in Filialdokumente (`_prg/2/*.doc`).
- Filialdokumente auf webverträgliche Namen umbenennen, Nummerierungen entfernen (`_prg/2a/*.doc`).

• Herstellen der Html-Version mit **Speichern unter...** (`_prg/2a/*.htm`). Dabei müssen unbedingt im rechten Frame die Gliederungsansicht eingeschaltet und die **Unterdokumente erweitert** sein.

• Links im Inhaltsverzeichnis und die Startdatei im Framedokuments müssen händisch korrigiert und als neue Dokumente `Inhalt.htm` und `default.htm` gespeichert werden.

Nachbearbeitung

Wenn das Originaldokument verändert wird, müssen danach folgende Operationen durchgeführt werden, um die aktualisierte Html-Version zu erhalten:

- Durchführen der Änderungen in den Doc-Dateien im Verzeichnis `_prg/2a`
- Löschen der Html-Filialdokumente (wichtig!)

Mit **Speichern unter...** die Html-Version herstellen (`_prg/2a/*.htm`). Dabei müssen unbedingt im rechten Frame die Gliederungsansicht eingeschaltet sein und die **Unterdokumente erweitert** sein

Zusammenfassung

Bei Ausnutzung der in Word verfügbaren Funktionen **Frames** und **Unterdokument** kann man mit einigen Tricks eine Menge Arbeit bei der Publikation eines Dokuments im Html-Format ersparen und erhält automatisch eine navigierbare und aufgeteilte Version des Originaldokuments.