

Größere Dokumente mit \LaTeX

Uwe Ziegenhagen

\LaTeX ist bekannt dafür, auch mit größeren Dokumenten jeder Art fertig zu werden, da die zugrundeliegende Architektur überaus stabil ist. Doch wie organisiert man ein Projekt mit hunderten von Seiten? Aus den Erfahrungen eines größeren Dissertationsprojekts möchte ich kurz berichten.

Dissertation & Werkskatalog

In den letzten Monaten hatte ich Gelegenheit, bei einer Dissertation in Kunstgeschichte \TeX nische Hilfe zu leisten. Der Umfang der Dissertation sollte sich als beeindruckend erweisen: 350 Seiten reiner Fließtext mit insgesamt 1800 Fußnoten, dazu ein Anhang mit Tabellen und ungefähr 150 Abbildungen, gefolgt von mehr als 50 Seiten Literaturverzeichnis. Das fertige PDF hatte eine Größe von ungefähr 210 MB.

Zusätzlich zur eigentlichen Dissertation war jedoch noch ein Werkskatalog zu erstellen, der fast 800 Abbildungen enthielt und am Ende eine Größe von mehr als einem Gigabyte haben sollte.

Organisation

Da mehrere hundert Kilometer zwischen Autorin und mir lagen, bot das Internet die einzig sinnvolle Möglichkeit, Dateien und Informationen auszutauschen. Alle Daten – abgesehen von den Bildern, die per DVD und gelber Post ausgetauscht wurden – lagen dabei in einem zentralen Subversion-Repository auf einem virtuellen Linux-Server. Neben Subversion gibt es noch eine Reihe anderer Versionsverwaltungssysteme wie Git und Mercurial, die für diesen Zweck sinnvoll sind, aufgrund persönlicher Erfahrungen und sehr guter Desktop-Integration über TortoiseSVN zog ich jedoch Subversion vor.

Für die Installation und Konfiguration von \TeX Live 2014 und passenden Editoren wurde Teamviewer genutzt. Teamviewer ist eine Anwendung, die es ermöglicht, einfach und bei gesicherter Verbindung Maus und Tastatur eines entfernten Rechners zu übernehmen. Teamviewer ist für private Anwender kostenlos und für nahezu alle Betriebssysteme verfügbar; daher eignet es sich sehr gut, um Konfiguration »vor Ort« vorzunehmen.

Die Verzeichnisstruktur war relativ überschaubar. Es gab einen Unterordner »Spielplatz«, in dem diejenigen Code-Schnipsel abgelegt wurden, die noch nicht ihren Weg in die Dissertation gefunden hatten. So ließen sich verschiedene Zitierstile, Schriftarten und ähnliches ausprobieren, ohne dass die Integrität der eigentlichen Dateien der Dissertation gefährdet wurde.

Ein weiterer Unterordner »Dissertation« enthielt dann die Dateien, die für die Übersetzung der eigentlichen Dokumente benötigt wurden. Es gab dort zwar den obligatorischen »Bilder«-Ordner, ansonsten lagen aber alle TeX-Dateien hier.

Aufteilung & Übersetzung des Dokuments

Die schiere Größe bedingt, die einzelnen Kapitel in separaten Dateien abzulegen und mittels `\include` einzubinden. Eine Fehlersuche wäre sonst schlichtweg nicht effizient durchführbar gewesen, da ein kompletter Durchlauf von `pdflatex` selbst auf einem schnellen Desktop-Rechner mit Intel i7-Prozessor, schneller SSD und 32 GB RAM bereits eine gute halbe Minute dauerte.

Da der genutzte TeXworks Editor entsprechende Unterstützung bietet, konnte in den einzelnen Kapitel-Dateien über `% !TeX root` festgelegt werden, welches die zu kompilierende Hauptdatei war, in die dann dieses entsprechende Kapitel inkludiert worden war (siehe Listing ??).

Listing 1: Definition der Hauptdatei in den einzelnen Kapitel-Dateien

```
% !TeX root = Dissertation.tex
...
\chapter{<Kapitelname>}
```

Für die eigentliche Steuerung des \LaTeX -Laufs wurde Arara genutzt. Dieses Programm, seit 2013 auch Teil von TeX Live, wertet spezielle Kommentare in der \LaTeX -Datei aus. So lässt sich beispielsweise steuern, dass zuerst zwei `pdflatex`-Läufe stattfinden müssen, bei denen kein PDF geschrieben wird, sondern nur die Hilfsdateien erzeugt werden. Im Anschluss kümmert sich `biber` um die Bibliografie, bevor zwei weitere `pdflatex` Läufe das fertige PDF setzen. Listing ?? zeigt einen entsprechenden Dokumentenkopf.

Listing 2: Arara-Kopf für die Steuerung des \LaTeX Laufs

```
%!TEX TS-program = Arara
% arara: pdflatex: { draft : yes }
% arara: pdflatex: { draft : yes }
% arara: biber
% arara: pdflatex
% arara: pdflatex
```

```
...  
\documentclass[11pt,bibliography=totoc,a4paper]{scrreprt}
```

Als kleiner Nachteil von Arara erwies sich, dass bei Fehlern während der Übersetzung gelegentlich Instanzen von pdf_latex im Speicher blieben, die dann manuell per Taskmanager geschlossen werden mussten.

Bibliografie

Die Bibliografie, mit knapp 800 Einträgen auch ein wenig größer, wurde mit JabRef 2.10 verwaltet. Diese kostenlose Software ist in Java geschrieben und bietet ein gut zugängliches grafisches Benutzerinterface für die manuell nur schwer handhabbaren Bib-Dateien. Zu beachten ist aber, dass JabRef standardmäßig den BibT_EX Modus ist und das ISO 8859 Encoding nutzt. Unter Optionen > Einstellungen > Allgemein lässt sich auf UTF-8 umschalten, unter Optionen > Einstellungen > Erweitert der BibL^AT_EX Modus aktivieren.

Bilder

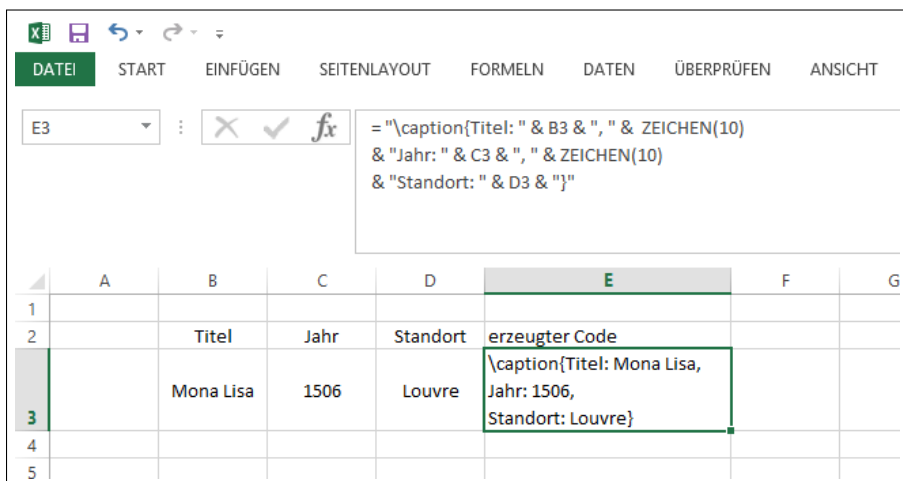
Da der Werkkatalog fast 800 Bilder enthielt, die jeweils 5–15 Zeilen Bildunterschrift bekommen sollten, wäre eine manuelle Bearbeitung der \caption{}-Befehle sehr mühselig und fehleranfällig gewesen. Daher wurde eine entsprechende Excel-Datei aufgesetzt, in deren fast zwanzig Spalten eine Vielzahl von Informationen, wie z. B. Katalognummer, Name der Bilddatei, Signatur, Maße, Provenienz, genutzte Technik und Literaturverweise zum jeweiligen Werk gespeichert wurden. Der eigentliche \caption Befehl wurde dann mit den Zeichenketten-Funktionen von Excel zusammengebaut. Das Stichwort ist hier Verkettten() bzw. die Abkürzung mittels &. Zeilenumbrüche im fertigen String können übrigens über Zeichen(10) erzeugt werden. Ein Beispiel für diese Vorgehensweise findet sich in Abbildung ??.

Die fertigen Bildunterschriften konnten dann einfach mittels Kopieren & Einfügen in das vorbereitete L^AT_EX-Dokument übernommen werden.

Querverweise

Eine weitere Besonderheit bei dieser Arbeit waren die zahlreichen Querverweise. Diese gab es nicht nur innerhalb von Dissertation und Werkskatalog, sondern auch zwischen beiden Dokumenten. Im Werkskatalog sollte nämlich auch aufgelistet werden, wo in der Dissertation dieses Werk erwähnt wird.

Was für normale Schreibprogramme wie Word oder OpenOffice ein Ding der Unmöglichkeit sein dürfte, war mit L^AT_EX recht einfach umzusetzen. Die Lösung fand

Abb. 1: Beispiel für durch Excel erzeugten \LaTeX -Code

sich in Heiko Oberdieks `zref-xr` und `zref-user` Paketen, die das Referenzieren von Labels aus anderen \TeX -Dateien erlauben. Ein vollständiges Beispiel habe ich unter <http://uweziegenhagen.de/?p=3020> abgelegt.

Erweiterung & Fazit

Aus \TeX nischer Sicht war diese Arbeit schon ein »ordentlicher Brocken«; neben dem Umfang von insgesamt 1300 Seiten waren es die vielen kleinen Anforderungen, die die Mitarbeit interessant werden ließen.

Es hat sich dabei aber auch gezeigt, dass man mittels \LaTeX auch komplexe Schriftstücke relativ bequem setzen lassen. Mehrere Sprachen wie Sütterlin, Russisch, Griechisch und normalem Deutsch in einem Dokument – kein Problem! Querverweise zwischen Dokumenten – kein Problem! Hunderte von Grafiken teilautomatisiert setzen – kein Problem!

Für mich hat \LaTeX mal wieder bewiesen, dass es noch längst nicht zum alten Eisen gehört. In diesem Sinne: Happy \TeX ing!