

Zur Entstehung des Datenbandes

Das weltweit akzeptierte Standardverfahren zur Speicherung genealogischer Daten ist das Format GEDCOM. Das ist die Abkürzung für **G**enealogical **D**ata **C**ommunication. Entwickelt wurde der GEDCOM-Standard von der Kirche Jesus Christus der Heiligen Letzten Tage, den Mormonen, die auch das Copyright besitzen.

GEDCOM ist ein Vorläufer für die modernen relationalen Datenbanken. 'Relational' bezeichnet hier den Sachverhalt, daß die Datenbank 'Beziehungen' speichert. Jede Person, die in der Datenbank verzeichnet wird, erhält eine Nummer, die sie eindeutig kennzeichnet (INDI - Nummer, von „Individual“). Nun spielt aber eine Person innerhalb einer genealogischen Datensammlung ganz verschiedene Rollen: So ist zum Beispiel ein Mann, der als „Stammvater“ einer Familie gilt, gleichzeitig auch das Kind eines Elternpaares, er hat vier Grosseltern, er ist selbst wiederum Vater, Grossvater, aber auch Bruder, Onkel oder Taufpate - kurz: Jede Person hat umfangreiche 'Beziehungen' zu anderen. Das gilt analog auch für Familien: Über die Personen, die zur Familie gehören, sind die Familien untereinander verbunden und miteinander verwoben .

Die genealogische Datensammlung der „Sehrbunds“ ist ebenfalls in einer GEDCOM Datei gespeichert. Die Datei enthält 82.694 Datenzeilen; darin speichert sie 3.842 Personen, davon 56 Prozent männliche. 873 verschiedene Geburtsnamen (Familiennamen) tauchen auf, zusätzlich 207 Personen, deren Namen noch nicht ermittelt werden konnten. Auf 308 Orte konzentriert sich die Herkunft; acht Orte davon versammeln fast die Hälfte aller Personen.

Nun stellte sich die Aufgabe, den Inhalt dieser umfangreichen Datensammlung in Buchform zu veröffentlichen. Doch wie können relationale Daten hintereinander auf Buchseiten dargestellt werden? Man könnte daran denken, alle Daten einer Person zusammenzufassen und die dann alle nacheinander abzudrucken. Aber in welcher Reihenfolge? Alphabetisch? Nach ihrem Auftreten im Zeitablauf? Und wie die Zusammenhänge und Verbindungen aufzeigen? Viele genealogischen Programme nutzen dafür die so genannte „Familienliste“, und auch wir haben uns dazu entschlossen. Dabei wird jede in der Datenbank als „Familie“ gekennzeichnete Gruppe von Menschen (jede mit einer **F**(amilien)-Nummer gekennzeichnet) im Zusammenhang dargestellt. Der Regelfall einer „Familie“ ist die Zusammensetzung Vater + Mutter + Kind(er), ergänzt durch Hinweise auf die Eltern von 'Vater' und 'Mutter', auf weitere Ehen, so sie denn bei 'Vater' oder 'Mutter' vorkommen, auf uneheliche Kinder, auf andere wichtige, gefundenen Informationen zu den Personen der Familie in Form von 'Anmerkungen' und 'Quellenhinweisen'. Die Quellen sind mit einer Nummer gekennzeichnet. Alle Quellen werden in einem eigenen Abschnitt im Volltext wiedergegeben. Alle Personen sind mit ihrer INDI-Nummer gekennzeichnet und sind damit in jedem Zusammenhang eindeutig zu identifizieren.

Wir haben die so zusammengestellten „Familiendatensätze“ nach dem Alphabet geordnet dargestellt, da dies die am meisten plausible und logische Form menschlicher Sucharbeit ist. Alphabetisch bedeutet hier: geordnet nach dem Familiennamen des Vaters der jeweiligen Familie. Zu dieser Sortierungsgruppe gehören die weitaus meisten Datensätze.

Da indes nicht alle Familien einen uns bekannten Vater haben, folgen dann die „Familien mit nicht bekannten Vätern“, alphabetisch geordnet nach dem Familiennamen der ‚Mutter‘. Dieser Gruppe folgt eine weitere, wo weder Vater noch Mutter bekannt sind, auch hier alphabetisch geordnet nach dem ersten sinngebenden Namen der Familie. Unbekannte Namen wurden - den Eingaben in der Datenbank folgend - als ‘?’ (Fragezeichen) oder „??“ (zwei Fragezeichen) oder als „unbekannt“ wiedergegeben.

Abgeschlossen wird die Darstellung mit einer Auflistung aller Informationen über solche Personen, die (noch) keiner Familie zugeordnet werden konnten. Auch die haben wir nach dem Alphabet sortiert.

Leider haben wir weltweit nicht eine einzige Genealogie-Software gefunden, die uns dabei helfen konnte, diese Zuordnungs- und Sortierungsarbeiten so vorzunehmen, wie es hier für die Herstellung eines Buches erforderlich ist. Selbst da, wo wir die Funktion „Druck von Familienlisten“ in einem Programm gefunden haben, konnten wir das nicht nutzen. Denn es galt ein Buch herzustellen, das nicht beliebig umfangreich sein durfte. Etwa 1.500 Familien auf 1.500 Seiten darzustellen - das hätte den Rahmen eines normalen, nutzbaren Buches gesprengt; es ist nur so möglich, indem die Familiendatensätze - optisch getrennt - laufend hintereinander gesetzt werden, anstatt für jeden Datensatz eine neue Seite zu beginnen.

So mußten wir einen Weg zur Buchherstellung beschreiten, der bisher so noch nie gegangen wurde. Mit einem eigens programmierten PERL-Script haben wir die Datenbank nach den Vorgaben des Autors ausgelesen, sortiert und dann in Form einer HTML-Seite ausgegeben. HTML-Seiten sind gewöhnliche Internet-Seiten. Wer sie kennt, weiss, daß eine solche Seite beliebig lang sein kann. Unsere bestand am Ende aus 108.919 Zeilen.

Diese HTML-Datei haben wir sodann in das „Portable Document Format“ (.pdf) umgewandelt; dabei haben wir das Buchseitenformat, die Seitenränder und die Nummerierung mit Seitenzahlen vorgegeben sowie die Druckauflösung in „Dot per Inch“ (dpi). Auf diesem Weg entstand ein .pdf - Dokument, welches den sehr speziellen Vorgaben der digitalen Buchherstellung entspricht und von dem aus sofort - ohne den Umweg über die Lithographie - gedruckt werden kann.

Mit diesem - einerseits eleganten - Weg haben wir monatelanges Abtippen und Korrekturlesen eingespart. Wir haben aber auch einige weniger elegante Schwächen damit akzeptiert. So sind die Seitenumbrüche nicht immer glücklich; zum Beispiel wird manchmal nur eine Zeile eines Familiendatensatzes auf die Folgeseite umbrochen, oder ein neuer Datensatz beginnt mit nur einer Zeile Inhalt am Fuß einer Seite. Diese kleinen Defizite ästhetischer und formaler Art bitten wir uns nachzusehen. Da die Datensammlung indes kaum der normalen Lektüre von Seite-zu-Seite dienen wird, sondern dem Finden und Nachspüren, halten wir es für vertretbar, einmal mehr umzublättern.

Richard Ebert
Hamburg, im Dezember 2005