

Kurzübersicht der Auswanderer-Homepage

von Tilo Wütherich

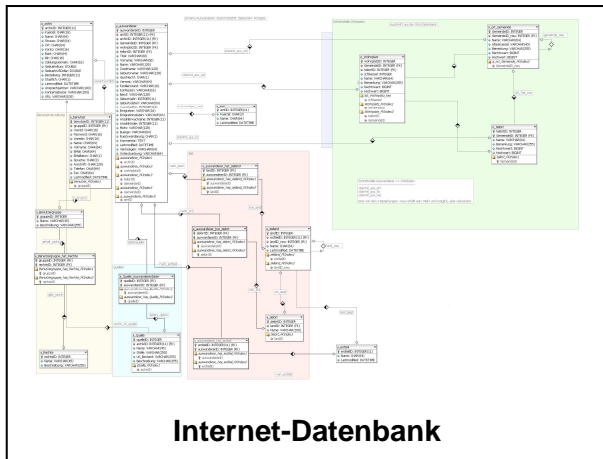
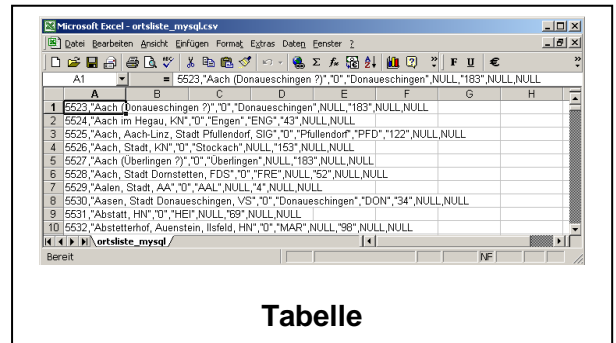
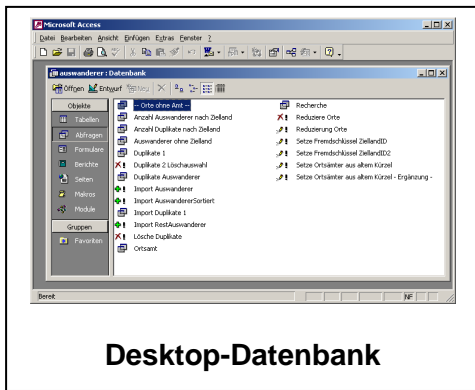
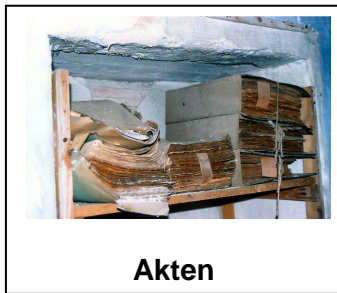
1	Projektgeschichte	1
2	Die gegenwärtige Technik	5
2.1	Datenbank und Pflegemodul.....	5
2.2	Wohnplatzdaten.....	5
2.3	SixCMS.....	6
2.4	MnogoSearch	6
3	URL's	6

1 Projektgeschichte

Die Auswandererdatenbank begann in den 70er Jahren mit der Übertragung ausgewählter Auswanderer-Daten von den Originalakten in mehrere Karteien (<http://de.wikipedia.org/wiki/Kartei>). Damit sollten Anfragen von Angehörigen oder Nachkommen von Auswanderern und Genealogen schneller beantwortet werden können. Mit dem Aufkommen der EDV in den Archiven Mitte der 80er Jahre wurde bald die elektronische Übertragung der Daten vorgenommen, einmal um den Sachbearbeitern die Auskünfte zu erleichtern, zum anderen sollten die Daten online im Lesesaal zugänglich gemacht werden.

Detailliertere Angaben zur Geschichte der elektronischen Datenbank fehlen, sie dürfte aber mehrfach zwischen Datenformaten konvertiert worden sein. Einen idealtypischen Werdegang zeigt Abbildung 1:

Abbildung 1: fiktiver Werdegang vom Archivgut zur Online-Anwendung



Dieser idealtypische Werdegang einer Online-Datenbank, der bei den Auswanderern aber nicht so abgelaufen ist, kann sich aus mangelnder Vorplanung, Zieldefinition und Informations-Fachwissen der Beteiligten ergeben. Bei der Auswanderer-Anwendung hat die Entwicklung der EDV in den letzten 30 Jahren die entscheidende Rolle gespielt. Mit dem Aufkommen der PCs wurde angefangen, die Karteidaten elektronisch zu erfassen. Mit der Verfügbarkeit leistungsfähigerer Software wurde das System dann schrittweise aus- und umgebaut.

Die Akten sind im Magazin nur sehr aufwendig zugänglich und nur für einen eng begrenzten Personenkreis. Die Karteien sind schneller einsehbar, enthalten aber nur einen Teil der Informationen und sind ebenfalls nicht für den Nutzer gedacht. Leider legen die Karteikarten ein scheinbar "natürliches", flaches Datenformat nahe. Sie bilden eine Ansammlung gleich strukturierter einzelner Datensätze, so daß es zunächst nahe liegt, die Daten in eine Tabelle zu übertragen, bei der jede Zeile einer Karteikarte entspricht. Die Karteikarten selbst können jedoch nach einem bestimmten System geordnet sein, z.B. nach Herkunftsort. Diese Anordnung würde in einer Datenbank durch die Verknüpfung (Relation) von Tabellen abgebildet, z.B. Wohnplatztabelle - Auswanderertabelle.

Schnell wird dann allerdings deutlich, daß viele Daten unnötig mehrfach, d.h. redundant, abgelegt sind. Im Fall der Auswanderer-Datenbank sind das z.B.

- Herkunftsorte: oft gibt es verschiedene Schreibweisen des gleichen Ortes ("Esslingen", "Eßlingen", "Esslingen a.N.", "Esslingen/Neckar" etc.), dagegen sind unterschiedliche Orte gleichen Namens nicht ohne weiteres trennbar (z.B. Dettingen, Möhringen, Altdorf...).
- Zielländer sind mehrfach redundant vorhanden, oft unsystematische Schreibweise, z.B. Nordamerika, USA, Pennsylvania, außerdem wird nicht zwischen Stadt - Staat -Land - Erdteil getrennt.
- zuständige Ämter und Archive
- Quellenangabe zur Recherche; bei Quellen in verschiedenen Archiven wurde der Auswanderer-Datensatz kurzerhand mehrfach angelegt, nur mit anderen Angaben im Quellen-Feld.

Das nächste Stadium ist dann eine Datenbank mit ad-hoc aufgesetzten Relationen, um die häufigsten Redundanzen zu vermeiden. In diese werden die vorhandenen elektronischen Daten übertragen, ggf. halbautomatisch. Leider ist die Datenkonsistenz bis dahin schon verloren gegangen. Oft wird das Problem ignoriert, weil die Anwendung in den meisten Fällen das erwünschte Ergebnis bringt. Abbildung 2 zeigt ein altes, relationales Schema auf Access-Basis. Auffallend ist, daß es 4 (!) Möglichkeiten gibt, ein Zielland einzutragen, in zwei verschiedene Volltextfelder und in zwei verschiedenen verknüpfte Tabellen.

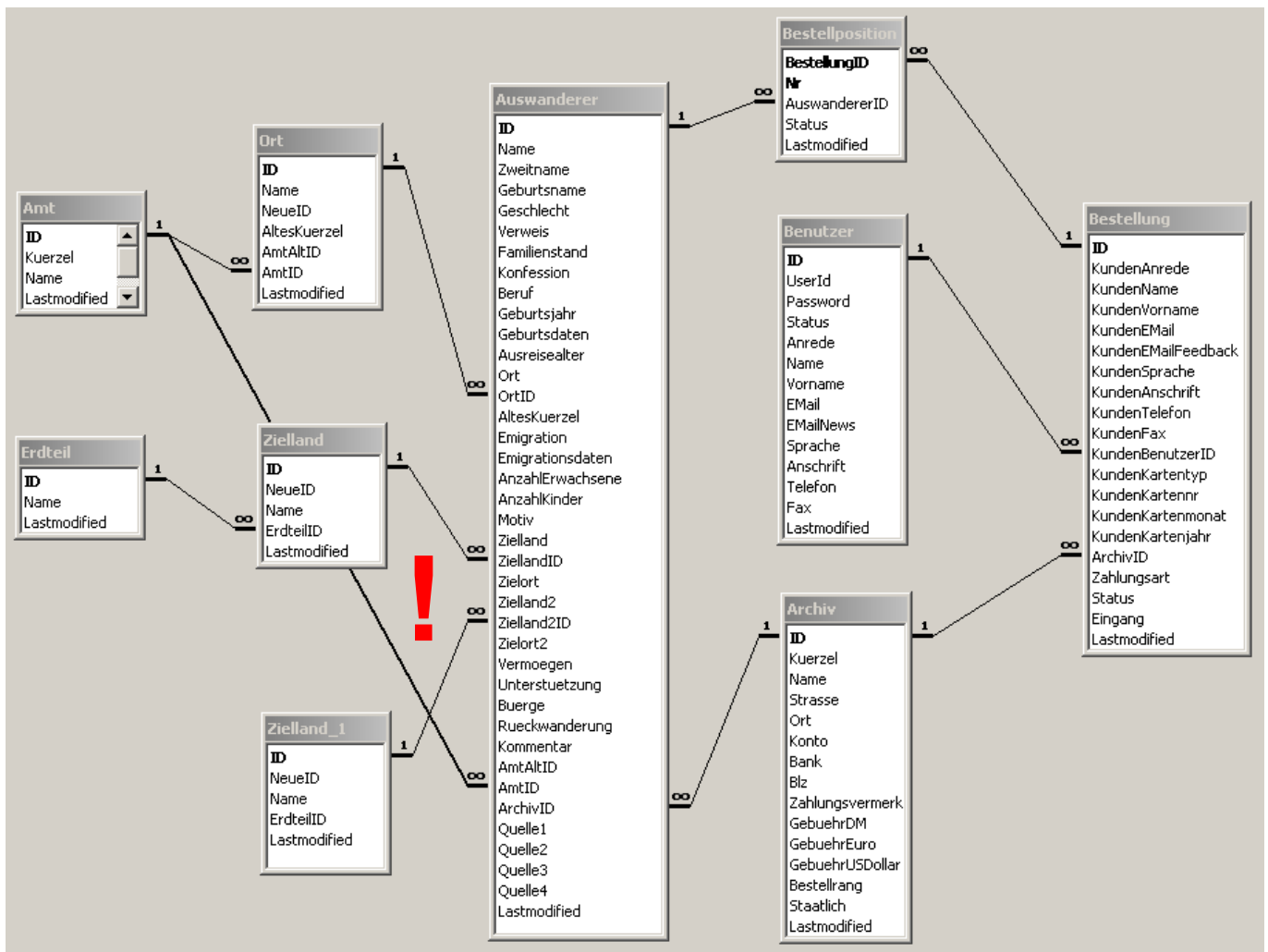


Abbildung 2: altes Access-Schema mit unsystematischen, mehrfachen Feldern, z.B. Zielland.

Im nächsten Schritt wird aus der inkonsistenten, mangelhaft gegliederten Desktop-Datenbank eine Online-Anwendung gemacht, mit erweiterten Recherchefunktionalitäten und Warenkorb. Spätestens hier ist aufgrund der neuen Funktionalitäten ein Redesign nötig, um z.B. angemeldete Benutzer, Warenkorb, Sachbearbeiter und Bestell-Workflow, Rechteverwaltung und Datenpflege umsetzen zu können. Spätestens hier macht sich auch die Inkonsistenz der Daten bemerkbar. Die Folge ist eine aufwendige "Bereinigung" der Daten, bei der z.B. verschiedene Ortschreibweisen vereinheitlicht werden etc. Leider kann dabei oft nicht mehr auf den Ausgangsdatenbestand, die Akten, zurückgegriffen werden, da es an Zeit und Geld bzw. Arbeitsleistung fehlt. Eventuell können sich deshalb neue Fehler bei der Datenbereinigung einschleichen.

Eine besondere Schwierigkeit bereitet die geographische Verortung der Auswandererdaten. Um Einzelheiten recherchieren zu können, muß bekannt sein, wo die Originalakten lagern. In den Staatsarchiven sind die Akten nach Bezirks- und Oberämtern abgelegt, deren Zuständigkeitsbereiche sich im Laufe der Jahrzehnte geändert haben. Man muß deshalb nicht nur den Herkunftsort eines Auswanderers kennen, sondern auch die damaligen Zuständigkeiten. Auch die Ortsnamen selbst und ihre politische bzw. verwaltungsrechtliche Zuordnung ändert sich im Laufe der Zeit. Deshalb ist es im Grunde nötig, eine zeitabhängige Verortung zu erstellen, oder eine auf kleinsten Raumeinheiten basierende, in unserem Fall der Wohnplatz mit eindeutiger ID und geographischen Koordinaten (RW/HW).

2.3 SixCMS

Dieses Content-Management-System beinhaltet die bis dahin als statische Html-Seiten vorhandenen Hintergrundinformationen, bietet Zusatzfunktionalitäten für den Benutzerkomfort (zumailen und ausdrucken einer Seite) und sorgt für ein einheitliches Layout. Dieser redaktionelle Teil kann ohne Programmierkenntnisse nach kurzer Einarbeitung von Mitarbeitern gepflegt werden. Ebenso bietet das SixCMS Funktionalitäten an, um mit relativ wenig Programmieraufwand ein Benutzerforum oder Serienbrief ("Newsletter") einzurichten.

Die Anwendung kann in verschiedenen Sprachen ausgegeben werden. Bisher sind Deutsch und Englisch umgesetzt. Die Artikel des SixCMS werden dazu in der entsprechenden Sprache erstellt und bekommen diese Eigenschaft zugeordnet (Sprachfeld de, en). Die Navigationsmenüs werden der Sprachauswahl und der Zuordnung der Artikel untereinander entsprechend automatisch aufgebaut, so daß nur deutsche oder nur englische Artikel ausgegeben werden. Alle sonstigen wörtlichen Ausgaben erfolgen über mehrdimensionale Variablen (Arrays), die je nach Sprachwahl durch das CMS eingebunden werden.

2.4 MnogoSearch

Für die Volltextrecherche in den mit SixCMS gepflegten Seiten wird die kostenlose Suchmaschine MnogoSearch benutzt. Diese ruft nacheinander alle Seiten auf und indiziert und gewichtet den Inhalt. Diese Metadaten werden wiederum in einer MySQL-DB hinterlegt, über die MnogoSearch dann eine Volltextsuche ermöglicht.

Somit ergibt sich für die Anwendung folgende Struktur:

3 URL's

- Startseite auf Pan: http://pan.bsz-bw.de/sixcms/detail.php?template=a_artikel&id=6441&sprache=de
- Startseite auf Auswanderer2: <http://auswanderer2.lad-bw.de/~moeller/auswanderer2/>
- PHPMyAdmin auf Auswanderer2: <http://auswanderer2.lad-bw.de/phpMyAdmin/index.php>

Abbildung 4: Schema Auswandereranwendung

