

APEX, SPSS und STATA

Klaus Hertle / Alexander Elsas
hertle.net / Goethe-Universität Frankfurt

Schlüsselworte:

Apex, SPSS, STATA

1 Einleitung

Das German Microdata Lab (GML) in Mannheim, Teil von GESIS, dem Leibniz-Institut für Sozialwissenschaften, stellt für interessierte Wissenschaftler Informationen aus den Mikrozensen zur Forschung und Auswertung über ihre Webseite zur Verfügung. Insbesondere die Downloadmöglichkeit von aufbereiteten Syntaxdateien für die Auswertung mit den Statistikprogrammen SPSS und STATA ist dabei eine Kernfunktionalität.

Realisiert ist dieser Service in Form einer APEX-basierten Backend-Anwendung zur Daten- und Sprachpflege und stellt somit die komfortable CMS-Lösung für die Frontend-Anwendung dar. Die ebenso APEX-basierte Frontendanwendung stellt dann die Informationen und Downloadmöglichkeiten zur Verfügung.

In diesem Beitrag wird die Entwicklung der Frontend- und Backendanwendung gezeigt und die damit gemachten Erfahrungen präsentiert und zur Diskussion gestellt.

2. Projekt

Interne Vorstudien auf Grundlage von einer MS-Access-Lösung führten zu der Entscheidung, insbesondere im Hinblick auf eine in der Projektfortschreibung zu erwartende öffentliche Zugänglichkeit, auf eine webbasierte Datenbank-Lösung zusetzen. Hier kam Oracle APEX in die engere Wahl. Die Vorteile von APEX, wie eine kurze Realisierungszeit und vorhandene Erfahrung in ähnlichen Projekten, einhergehend mit geringen Entwicklungskosten, gaben dabei den Ausschlag. Als weiterer Vorteil für ein APEX-System, stellte sich die Systemunabhängigkeit dar, da ein Wechsel der Server-Plattformen von Linux zu Windows als wahrscheinlich galt.

Im Jahr 2007 entstand auf Grundlage von APEX 2.1 eine Vorversion für die internen Forschungsk Kooperationen mit Projektpartnern.

Auf Grundlage der dabei gewonnenen positiven Erfahrungen wurde ab dem Jahr 2010 die Erweiterung um Mehrsprachigkeit und Vorbereitung des öffentlichen Betriebs vorangetrieben. Folgerichtig wurde die System-Migration auf APEX 4 durchgeführt und so eine nahtlose Integration in den vorhandenen Webauftritt (Design und Funktionalität) und die Netzwerkinfrastruktur ermöglicht.

3 Anwendungsumsetzung

3.1 Konzept

Grundlegendes Konzept der Anwendung ist zunächst die Aufteilung in:

- eine Backend-Anwendung,
- eine Frontend-Anwendung in Form der öffentlich verfügbaren Website und
- einen externen Programmteil zur Skriptgenerierung.

Frontend und Backend befinden sich in einem gemeinsamen APEX-Workspace.

3.2 Backend

Das Backend ermöglicht über einen passwortgeschützten Zugang die Administration der Anwendung und Pflege der Inhalte.

Im Datenbestand werden Mikrozensus auf Erhebungsbasis verwaltet (aktuell zwischen 1962 - 2006), eine Ergänzung weiterer Jahrgänge und Bearbeitung bereits eingepflegter Datenbestände ist jederzeit möglich. Jahrgänge können dabei einzeln im Frontend freigeschaltet oder ausgeblendet werden. Dies erlaubt die interne Bearbeitung von (neuen) Zensusjahrgängen.

Edit	Thema	Pos.	Variable	Variablenlabel	Ausprägungen	Verfügbarkeit	Ausgangsv.	Bemerkung
	Bildung	501	kiga	Kindergarten: gegenwärtiger Besuch	Auspr.	verfügbar	Ausgangs.	
	Bildung	502	sbes	Schule: gegenwärtiger Besuch	Auspr.	verfügbar	Ausgangs.	
	Bildung	503	schul_1	Art der gegenwärtig besuchten Schule: Schultyp	Auspr.	verfügbar	Ausgangs.	
	Bildung	504	schul_2	Art der gegenwärtig besuchten Schule: typisiert	Auspr.	verfügbar	Ausgangs.	
	Bildung	505	schul_3	Art der gegenwärtig besuchten Schule: Klassenstufe	Auspr.	verfügbar	Ausgangs.	
	Bildung	506	sab	Vorhandensein eines allgemein bildenden Schulabschlusses	Auspr.	verfügbar	Ausgangs.	
	Bildung	507	asa_1	Allgemein bildender Schulabschluss	Auspr.	verfügbar	Ausgangs.	
	Bildung	508	asa_2	Allgemein bildender Schulabschluss: detailliert	Auspr.	verfügbar	Ausgangs.	
	Bildung	509	berab	Vorhandensein eines berufsbildenden Abschlusses	Auspr.	verfügbar	Ausgangs.	
	Bildung	510	bab_1	Berufsbildender Abschluss	Auspr.	verfügbar	Ausgangs.	

Hinweis

- Neue Variable anlegen: Button "Neu"
- Eintrag bearbeiten: Symbol in der Spalte "Edit"
- Variablenübersicht: Klick auf Variablenname.
- Verfügbarkeiten anzeigen / bearbeiten: Klick auf verfügbar.
- Ausprägungen anzeigen / bearbeiten: Klick auf Auspr.
- Ausgangsvariablen anzeigen / bearbeiten: Klick auf Ausgangsv.
- Themenbereich wechseln: Auswahlliste "Thema"

Abb. 1: Backend - Zielvariablen bearbeiten.

Dateien (wie Handbücher und Dokumentationen) können in die APEX-Umgebung hochgeladen und später über das Frontend zum Download angeboten werden.

Die Anwendung wird im internationalen Umfeld mehrsprachig angeboten, Übersetzung und Sprachversionsverwaltung finden im Backend statt. Zur Zeit sind Deutsch und Englisch realisiert, eine Ausweitung ist vorbereitet. Die Implementierung der Mehrsprachigkeit erfolgt wie in (Alexander Elsas: CoBox - Mehrsprachige Apex-Anwendung zur Mitarbeiterkommunikation, DOAG 2011; verfügbar unter <http://databaselab.org/doag/cobox.pdf>) dargestellt.

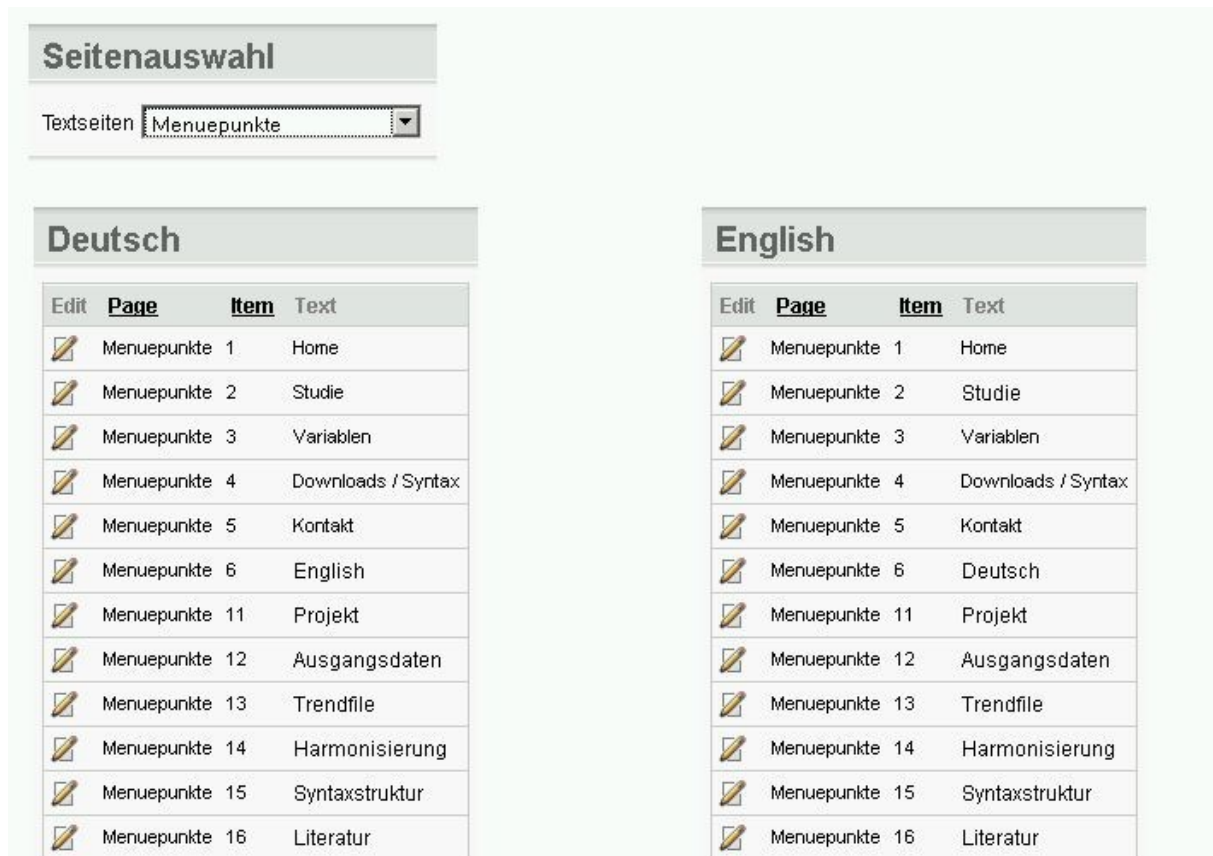


Abb. 2: Backend - Internationalisierung

3.3. Frontend

Das Frontend erlaubt den Nutzern passwortfreie Datenrecherche, Informationsabruf und den Download von individuell generierten Syntaxdateien für statistische Programmpakete. SPSS und STATA sind bereits integriert, eine Erweiterung um weitere Programmpakete, wie z. B. SAS, ist vorgesehen.

Dem Benutzer steht eine Sprachauswahl zur Verfügung, aktuell sind über das Backend Deutsch und Englisch angelegt, das Frontend ist für weitere Sprachen vorbereitet.

Im Frontend wird dem Nutzer ein Bestellformular präsentiert, das durch Anklicken die Auswahl der gewünschten Themen, Jahre und des genutzten Statistik-Paketes ermöglicht. Aus der Datenbank werden dann automatisch die benötigten Syntax-Job-Routinen generiert, die zur Erstellung eines kumulierten Datensatzes benötigt werden. Dabei werden fachliche Voraussetzungen berücksichtigt. Diese Syntax-Jobs werden dann als Downloadpaket inklusiv einer Dokumentation angeboten.

Abb. 3: Frontend - Syntax-Bestellformular (Ausschnitt)

3.4. Externer Programmtteil

Die vom Benutzer individuell angeforderten Syntax-Job-Routinen werden über ein externes, selbst entwickeltes Perl-Skript generiert. Die Benutzerwünsche werden aus dem Bestellformular über das APEX-System an das Skript weitergegeben.

Das Skript initiiert eine Datenbankverbindung und ruft alle benötigten Informationen ab. Es generiert anschließend die syntaxkorrekten Jobroutinen für unterschiedliche Statistikpakete und stellt alle erforderlichen Einzeldateien zusammen. Ein daraus gepacktes zip-Archiv wird dann an die Nutzerschnittstelle zum Download durchgereicht

Kontaktadressen

Klaus Hertle
 hertle.net
 Dammstraße 7
 D-68169 Mannheim

Telefon: +49 (0) 621-319 69 22
 Fax: +49 (0) 621- 454 91 96
 E-Mail: hertle@hertle.net
 Internet: hertle.net

Alexander Elsas
 Goethe-Universität
 Grüneburgplatz 1
 D-60323 Frankfurt

Telefon: +49 (0) 69-798 33636
 Fax: +49 (0) 69-798 33639
 E-Mail: aelsas@finance.uni-frankfurt.de
 Internet: aelsas.de