

## Internationale Studie – Probleme und Lösungen bei der Programmierung eines elektronischen Interviews

Lübke A, Reineke A, Pelz I

Bremer Institut für Präventionsforschung und Sozialmedizin (BIPS), Bremen

luebke@bips.uni-bremen.de

### Einleitung und Problembeschreibung

Im Rahmen der multizentrischen internationalen epidemiologischen Studie BiCoS („Bitumen Cohort Study“: Kohortenstudie der Krebssterblichkeit unter Asphaltarbeitern) wurde eine Datenbank entwickelt, die von den Zentren der beteiligten Länder zur Datenerhebung eingesetzt wird. Koordiniert wird die Studie von der IARC (International Agency for Research on Cancer, Lyon). Zu den teilnehmenden Ländern gehören Dänemark, Finnland, Frankreich, Deutschland, Israel, Norwegen und die Niederlande.

Da nicht in allen Zentren eine Datenbankprogrammierung realisierbar war, wurde die Programmierung zentral für alle Studienpartner vom BIPS durchgeführt. Der große Vorteil die Daten bei der Durchführung einer Studie in allen Zentren in einer vorgegebenen Datenbank aufzunehmen liegt darin, dass jede Einrichtung die gleichen Datenstrukturen und Dateiformate benutzt. Das erleichtert dem Datenzentrum das Zusammenspielen der Daten, reduziert die Fehlerrate und bedeutet eine beachtliche Zeitersparnis bei der Datenaufbereitung des Gesamtdatensatzes. Außerdem wurde die Datenbank mit Plausibilitäts-, Konsistenz- und Vollständigkeitskontrollen programmiert, sodass die Daten in allen Einrichtungen möglichst korrekt und präzise aufgenommen werden können.

Damit alle Partnerländer diese Datenbank benutzen können, mussten mehrere Problemfragen/-situationen gelöst werden, u.a.:

- Welches Datenbankmanagementsystem ist bei der Durchführung der Studie am sinnvollsten?
- Als Datenbankgrundlage galten Fragebögen in englischer Sprache, die gemeinsam von den beteiligten Instituten entwickelt wurden. Die Datenbank musste so konstruiert werden, dass es den Studienpartnern möglich ist, die Texte, die auf der Benutzeroberfläche abgebildet werden, einfach und sicher in die landeseigene Sprache zu übersetzen.
- In manchen Texten/Fragen muss der Nachname des an Krebs verstorbenen Asphaltarbeiters eingeblendet werden (Interviewpartner sind Angehörige, Bekannte und Kollegen, die zu dem verstorbenen Asphaltarbeiter befragt werden).
- Die Datenbank musste so entwickelt werden, dass die Durchführung von Updates in den Studienzentren vor Ort möglich ist.

### Lösungsansatz

Damit für alle Studienpartner eine kostengünstige Lösung geschaffen werden konnte, wurde als Datenbankmanagementsystem MS Access gewählt. Dieses hatte den Vorteil, dass große Ausgaben für die Beschaffung anderer Software ausblieben, da fast alle Studienzentren über MS Access verfügten. Die Datenerhebung erfolgt in den Studienzentren vor Ort am PC als Telefon- bzw. als ‚face to face‘ Interview mit Hilfe von Laptops. Deshalb war eine einfache Installation und Einrichtung des Datenbankmanagementsystems, welches bei MS Access gegeben ist, mit ausschlaggebend.

Das Datenbanksystem der Studie besteht aus zwei Datenbankdateien: die erste enthält die Formulare, Module, Abfragen usw. (mde-Datei) und die zweite enthält alle Tabellen, die über eine Verknüpfung mit der Formulardatenbank in Beziehung stehen (mdb-Datei). So können Änderungen an den Formularen oder an den Tabellen gesondert behandelt werden. Wenn sich das Update nur auf Objekte beschränkt, die sich in der Formulardatenbank befinden, kann die in den Instituten vorhandene Formulardatenbank einfach durch die aktualisierte ersetzt werden, ohne dass Daten verloren gehen. Modifikationen an den Tabellen und Datenstrukturen werden mit VBA-Prozeduren ausgeführt, die die gewünschten Änderungen erzeugen. Die im BIPS entwickelten und getesteten VBA-Prozeduren werden den teilnehmenden Studienpartnern in einer separaten Datenbank zur Verfügung gestellt. Diese können dann per Schaltflächenbetätigung die Prozedur ausführen und die Modifikationen an den Tabellen werden durchgeführt. Durch dieses Verfahren ist gewährleistet, dass die bereits erhobenen Daten in den Tabellen erhalten bleiben. Die Update-Datenbank ist so konstruiert, dass es den Anwendern nicht möglich ist, ein Update auszulassen oder es mehrmals auszuführen.

Vorteil der eingesetzten mde-Datei ist nicht nur die Verringerung der Datenbankgröße und die daraus resultierende Verbesserung der Systemleistung, sondern es ist vor allem ein Schutz vor unberechtigter Veränderung der Formulare und Module, da der Quellcode nicht eingesehen werden kann. Andererseits können aber auch die Texte, die über die Bezeichnungsfelder der Formulare angezeigt werden, nicht geändert bzw. in die landeseigene Sprache übersetzt werden. Um dieses Problem zu lösen wurde ein VBA Modul entwickelt und implementiert, welches die Texte, die auf der Benutzeroberfläche angezeigt werden sollen, aus einer Tabelle einliest und auf dem aktiven Formular einblendet. Dazu werden alle Texte die in dem Interview vorkommen und in die landeseigene Sprache übersetzt werden sollen, in einer Tabelle gespeichert. Abbildung 1 zeigt einen Ausschnitt aus dieser Tabelle.

qname	english	country
N02Q00T03	First, I would like to ask you about your relationship to Mr. \$\$\$ and when you got to know him:	Zu Beginn möchten wir Sie fragen, in welcher Beziehung Sie zu Herrn \$\$\$ standen und seit wann Sie ihn kannten.
N02Q04T01	4. What relation was Mr. \$\$\$ to you?	4. In welchem Verwandtschaftsverhältnis / welcher Beziehung standen Sie zu Herrn \$\$\$?

Abb. 1: Datenbanktabelle

Die Schlüsselvariable *qname* enthält die eindeutigen Bezeichnungsfeldnamen der Formulare. In der Spalte *english* wurden die englischen Texte von den Papierfragebögen eingetragen. Dabei darf der Text in einer Zelle maximal 256 Zeichen lang sein. Die Spalte *country* enthält den übersetzten Text in der landeseigenen Sprache. Die Übersetzung findet nicht in der Tabelle direkt statt, sondern aus Sicherheitsgründen und zur Unterstützung der Benutzerfreundlichkeit über ein Formular. Über das Formular werden die Spalten aus Abb. 1 angezeigt, wobei *qname* und *english* gesperrt sind, sodass hier keine Änderungen an den Daten vorgenommen werden können. Die Übersetzung kann einfach über das Formular in die Datenfelder der Spalte *country* eingetragen werden.

Mit Hilfe des bereits erwähnten VBA-Moduls werden die Texte aus der Spalte *country* eingelesen und über die Bezeichnungsfelder auf den Formularen angezeigt.

Dem Modul werden die Öffnungsargumente und der Formularname des aktiven Formulars als Parameterwerte übergeben. Anschließend werden alle Steuerelemente des aktiven Formulars durchlaufen und überprüft. Liegt als Steuerelementtyp ein Bezeichnungsfeld vor, findet ein Abgleich des eindeutigen Bezeichnungsfeldnamens vom Formular mit der Tabelle (Spalte *qname*) statt und der zum Datensatz gehörige übersetzte Text aus der Spalte *country* wird selektiert. Danach wird der selektierte Text auf einen vorgegebenen String "\$\$\$" untersucht. Diese Zeichenfolge dient als Platzhalter für den Nachnamen der Indexperson, der manchmal innerhalb der Texte angezeigt wird. Anstelle des Platzhalters wird durch eine Verkettung der Nachname des Asphaltarbeiters eingebunden, der in jedem Formulkopf eingeblendet wird. Anschließend wird dem Bezeichnungsfeld der selektierte Text inklusiv des ausgetauschten Platzhalters zugewiesen und auf dem Formular in der landeseigenen Sprache angezeigt.

### Möglichkeiten und Grenzen

Das Einlesen und Anzeigen der Texte aus den Tabellen über die VBA-Module funktioniert sehr gut. Es gibt keine Verzögerungen beim Anzeigen der Formulare. Ein Nachteil ist jedoch, dass die Größe der Bezeichnungsfelder festgelegt werden muss. Das bedeutet, dass die Texte an die Bezeichnungsfelder angepasst werden müssen. Ist der einzublendende Text viel kürzer, als die festgelegte Größe des Bezeichnungsfeldes, müssen Kompromisse beim Formular-Layout eingegangen werden.

Analog zur dargestellten Methode, die übersetzten Texte für die Bezeichnungsfelder der Formulare aus einer Tabelle zu selektieren, werden die Informationen zur Erstellung von „Messageboxen“ einer Tabelle entnommen und über ein Modul nach dem Aufruf in der VBA-Programmierung als „Messagebox“ ausgegeben. Somit können auch die Mitteilungstexte in die landeseigene Sprache übersetzt werden.

Zusätzlich können in der Datenbank alle Einträge aus den Nachschlagelisten der Kombinationsfelder über Formulare ggf. sortiert, erweitert, übersetzt und somit in der landeseigenen Sprache angezeigt werden. Dazu müssen die Werte, die in den Nachschlagelisten aufgelistet werden sollen, in Tabellen gespeichert werden.

Anfängliche Probleme mit der Datenbank im israelischen Studienzentrum konnten behoben werden. Die Ländereinstellungen stehen dort auf Hebräisch, was nicht mit Umlauten und Sonderzeichen harmonisierte, die in der Datenbank benutzt wurden. Dabei ist zu beachten, dass bei der deutschen Version von MS Access automatisch von der Software an verschiedenen Stellen Sonderzeichen in der Namenvergebung verwendet werden, die anschließend umgeändert werden müssen. Auch konnten Passwörter nur aus Zahlen konstruiert werden, da ein Abgleich der hebräischen Zeichen mit unserem Alphabet fehlschlug.

#### **Fazit**

Durch die VBA-Module für die Texteinblendung konnten Zeit und Kosten eingespart werden, die entstanden wären, wenn sich jedes Studienzentrum um eine funktionsfähige Datenbank in der landeseigenen Sprache hätte kümmern müssen. Durch das Benutzen einer gemeinsamen Datenbank wird das Ziel einer hohen Datenqualität erreicht und dem Datenzentrum bleibt viel Arbeit und Zeit bei der Erstellung des Gesamtdatensatzes und der Datenaufbereitung erspart. Eingabefehler und Inkonsistenzen in den Daten werden durch programmierte Plausibilitäts- und Vollständigkeitskontrollen reduziert. Im Fall der BiCoS Studie ist MS Access eine geeignete Software. Vor allem wegen der einfachen Handhabung bei der Installation ist es vorteilhaft, da die Studienzentren die ‚face to face‘ Interviews per Laptop durchführen. Aber es wäre von Vorteil gewesen, für die Datenbankprogrammierung eine englische Accessversion zu verwenden. Dadurch wären die Probleme mit der hebräischen Ländereinstellung nicht aufgetreten.