



Projektbeschreibung

QDDS-2: Software-Werkzeuge zur Dokumentation der Fragebogenentwicklung.

Rainer Schnell

1. Dezember 2005

Center for Quantitative Methods and Survey Research
Universität Konstanz

1 Projektziel

Bei methodisch anspruchsvollen Studien wird ein Fragebogen Dutzenden von Revisionen unterworfen. Die Revisionen sollten sowohl für die Datenanalyse und Datenarchivierung als auch im Hinblick auf die Weiterentwicklung in späteren Surveys in jedem Detail dokumentiert werden. Diese ausführliche Form der Dokumentation ist ebenso für die Lehre in der Umfrageforschung unverzichtbar. Trotzdem erfolgt die Dokumentation der Entwicklung eines Fragebogens üblicherweise aber nur partiell oder gar nicht. Eine der Ursachen hierfür liegt im Mangel an geeigneten Dokumentationswerkzeugen. Ziel des Projekts ist die Neuentwicklung eines Programm-Systems, das die dauerhafte elektronische Dokumentation der Fragebogenentwicklung und des Endzustandes des Instrumentes erlaubt.

2 Zusammenfassung des Forschungsstandes

Für die Dokumentation der Fragebogenentwicklung leisten bestehende Programme wie RCS (Revision Control System) oder QDDS (Questionnaire Development Documentation Script) nicht das Gewünschte: RCS erfordern für jeden Entwicklungsschritt das Speichern der jeweiligen Fragebogenversion als Ganzes und erlaubt faktisch kaum eine Dokumentation der Ursachen der Änderungen; QDDS erfordert eine manuelle Erstellung der Indexverwaltung und eine sehr aufwändige Produktion der Frageimages. Zusammenfassend kann festgehalten werden, dass die Dokumentation der Änderungshistorie von Fragebögen noch nicht in adäquatem Maße durch Softwaresysteme unterstützt wird.

2.1 Eigene Vorarbeiten/Arbeitsbericht

Für die Durchführung des Projekts kann auf umfangreiche Vorarbeiten zurückgegriffen werden.

2.1.1 Durchführung kognitionspsychologischer Pretests zur Instrumentenentwicklung

Die Arbeitsgruppe Konstanz hat im Rahmen von Drittmittel- und Lehrforschungsprojekten seit 1997 mehrere, sehr detaillierte Instrumentenkonstruktionen durch Anwendung kognitiver und traditioneller Pretest-Techniken durchgeführt. Dazu gehörten exemplarische Instrumente für Bürgerbefragungen (Waber 1999), Bibliothekbefragungen (Kreuter/Schnell 2000), Absolventenbefragungen (Kreuter/ Schnell/Kopp 2000) und Kriminalitätsfurchtsurveys (Kreuter 2002, Coutts 2002).

Die dabei entstandenen Instrumente wurden von mehreren Arbeitsgruppen aufgegriffen, so z. B. bei der Befragung der Benutzer der Universitätsbibliotheken in NRW durch Infas sowie einer neueren Kriminalitätsstudie in Hamburg, die ebenfalls durch Infas durchgeführt wurde.

2.1.2 Entwicklung eines Prototypen: QDDS

Zur Dokumentation der Fragebogenentwicklung des DEFECT-Projekts wurde ein neues Dokumentationssystem entwickelt: QDDS (Schnell/ Kreuter 2000, Schnell/ Kreuter

2002)¹. Bei QDDS handelt es sich im Wesentlichen um ein weniger als 350 Zeilen umfassendes Programm in der Programmiersprache Perl. QDDS ist ein Perl-Skript, das aus getrennt gespeicherten Teilen eines Fragebogens je nach interaktiver Benutzeranforderung in einem normalen Web-Browser eine Ausgabeseite erzeugt. Es basiert also auf der dynamischen Erzeugung von HTML-Seiten zur Fragebogendokumentation. Zur Dokumentation der Entwicklung eines Fragebogens mit QDDS werden zwei verschiedene Eingabedateien verwendet: eine Indexdatei (ASCII-File) und zahlreiche Frageimages (GIF-Dateien). Die Indexdatei enthält die Daten einer Matrix, deren Spalten die Versionen des Fragebogens darstellen und deren Zeilen die Kennnummern der Fragen enthalten. Jede Zeile enthält die Entwicklungsgeschichte einer Frage und jede Zelle enthält eine Fragenummer. Diese Nummer entspricht der Nummer der entsprechenden Frage in der jeweiligen Version des Fragebogens. Die Frageimagefiles sind Graphikfiles, die jeweils das Bild einer Frage enthalten. Je nach Art der Benutzerabfrage wird durch das QDDS eine neue HTML-Seite generiert, die entweder eine vollständige Version des Fragebogens wiedergibt (Fragebogenansicht) oder die vollständige Entwicklung einer Frage darstellt (Frageansicht). Zur Nutzung von QDDS benötigt man einen Perl-Interpreter und das von Schnell/Kreuter entwickelte Perl-CGI-Skript („qdds.pl“). Um QDDS nutzen zu können, muss man nicht Perl beherrschen, sondern nur die Konfigurationsdatei „config.txt“ mit einem normalen Editor bearbeiten. Des Weiteren benötigt man eine Bildverzeichnisdatei, die die Bilder der einzelnen Fragebogenversionen und Frageversionen beinhaltet. Diese Frageimages müssen in dem in der Konfigurationsdatei angegebenen Bildformat (z.B. GIF oder JPG) gespeichert werden. Für den Text am Anfang der Fragebogenansicht benötigt man die Versionskopfddatei und für den Text am Anfang der Frageansicht die Fragenkopfddatei. Die für dieses Programm wesentlichsten Dateien sind die Indexdatei und die Dokumentationsdatei.

Aus Gründen der Übersichtlichkeit empfiehlt es sich die Erstellung und Verwaltung der Indexdatei mit einem Spreadsheet-Programm (z.B. Excel) durchzuführen. Bei der Dokumentation werden die Fragen in der Abfolge der ersten Fragebogenversion in die erste Spalte eingetragen. Für jede weitere Version des Fragebogens wird für jede Frage in die nächste Spalte der Indexdatei die Nummer eingetragen, die die Frage in dieser Version des Fragebogens hat. Zusätzlich wird die Zelle, in der in der vorherigen Version etwas verändert wurde mit einem Stern gekennzeichnet. Wird eine Frage in einer späteren Version in zwei Teilfragen aufgeteilt, so wird dies in der Zelle durch ein Plus-Zeichen zwischen den beiden Fragennummern der Teilfragen indiziert. Werden zwei Fragen zusammengefasst, wird in beiden Zeilen die Nummer der neuen Fragen eingetragen. Wird eine Frage gelöscht, so bleiben die Zellen in den folgenden Spalten leer. Wird eine neue Frage in den Fragebogen aufgenommen, so wird sie in der Indextabelle unterhalb der alten Fragen eingefügt. Die Zellen der alten Fragebogenversionen bleiben für diese Frage leer. Die Dokumentationsdatei enthält für jede Frage einen Absatz, der die inhaltlichen Gründe für die Veränderung der Frage dokumentieren soll.

QDDS wurde auf der QDET („International Conference on Questionnaire Development,

¹ Die Implementierung kann im Internet interaktiv verwendet werden; sie findet sich unter <http://esem.bsz-bw.de/sicher>. Die Software kann unter <http://www.uni-konstanz.de/FuF/Verwiss/Schnell/qdds.zip> heruntergeladen werden.

Evaluation, and Testing Methods“) am 14-17 November 2002 in Charleston (South Carolina) vorgestellt (Schnell/Kreuter 2002).

2.2 Entwicklung von Softwaresystemen für die Sozialwissenschaften

Die Abteilung für informationswissenschaftliche und -technische Forschung und Entwicklung am IZ Sozialwissenschaften führt serviceorientierte Projekte zur Verbesserung der hauseigenen Produkte und Dienstleistungen und zur Unterstützung der sozialwissenschaftlichen Zielgruppen durch. Neben diesen Arbeiten im Schwerpunkt „Information Retrieval“ und Softwareergonomie im Kontext von integrierten Informationssystemen werden dabei auch Systeme zur Qualitäts- und Effizienzsteigerung von spezifischen Prozessen in den Sozialwissenschaften entwickelt.

Zur Unterstützung der redaktionellen Abläufe und des gesamten Produktionsprozesses der Reihe „Sozialwissenschaftlicher Forschungsinformationsdienst“ (soFid), der in je 28 thematischen Einzelbänden zweimal pro Jahr vom IZ als Printprodukt und interaktive CD-ROM produziert wird, wurde eine Lösung zum datenbankbasierten Publizieren geschaffen. Sie basiert auf mehreren ineinander greifenden Werkzeugen für spezifische Aufgabenstellungen, von der Recherche neuer Dokumente in den Datenbanken des IZ über die redaktionelle Selektion und Zuordnung dieser Dokumente zu den Kapitelgliederungen der 28 Themenbände bis zur Erzeugung der qualitativ hochwertigen Druckvorstufe und der Ausgangsdatei für die CD-ROM-Produktion. Hierbei kommen eine Reihe unterschiedlicher Technologien und Standards wie Java, XML und XML Stylesheet Transformationen (XSLT) zum Einsatz. Inhalte und Struktur jedes Einzelbands wird in einer Datenbank dokumentiert, so dass zu jedem Zeitpunkt ältere Ausgaben originalgetreu wieder erzeugt werden können.

Für das International Social Survey Programme (www.issp.org) hat das IZ unter inhaltlicher Beratung durch das Zentralarchiv für Empirische Sozialforschung (ZA) in Köln den ISSP DataWizard entwickelt. Hierbei handelt es sich um eine Software zur Homogenisierung länderspezifischer Umfragedaten auf eine gemeinsame semantische und formale Struktur (Standard Setup) hin. Der ISSP DataWizard erlaubt die Extraktion einzelner Variablen aus den in den 38 ISSP Partnerländern durchgeführten nationalen Erhebungen (z. B. dem ALLBUS in Deutschland) und die Recodierung ihrer Wertebereiche auf einen gemeinsamen Wertebereich (z. B. landesspezifische Parteien auf ein link-mitte-rechts Spektrum). Der Transformationsprozess wird durch unterschiedliche Visualisierungen unterstützt, durch Prüf- und Änderungsregeln automatisiert und vollständig dokumentiert. Die Software wurde plattformneutral in Java programmiert und nutzt relationale Datenbanken zur Datenspeicherung.

3 Ziele und Arbeitsprogramm

3.1 Ziele

Ziel des Projekts ist die Neuentwicklung eines Programms zur Dokumentation der Fragebogenentwicklung. Neben der Dokumentation einzelner Frageformulierungen und Formatierungen soll das Programm unaufwändig ein optimales Layout und eine hohe Druckqualität für Paper&Pencil-Surveys garantieren. Schließlich soll das Programm die Dokumentation vorhandener Fragebögen ermöglichen, in dem formale Fragekriterien wie Fragetyp, Art der Antwortvorgaben, Fragegegenstand und Kontext der Frage (Filterführungen, Stellung im Fragebogen) gespeichert werden¹.

Ein Fragebogen besteht aus einer definierten Abfolge einzelner Fragen, die sich im Laufe der Bearbeitung des Fragebogens ändern kann. Eine Frage besteht aus einer variablen Anzahl einer endlichen Menge von Elementen (Nummer, Text der Frage, Formularelemente, Verweise auf andere Fragen usw.). Diese Formularelemente (z.B. Checkboxes für Multiple-Choice-Fragen, Leerfelder für Eintragungen des Befragten) unterliegen Regeln für die Anwendung (z.B. immer mindestens 2 Checkboxes) und das Layout (z.B. erst Label, dann Checkbox; Ausrichtung rechtsbündig). Alle Änderungen an den genannten Bestandteilen eines Fragebogens müssen in ihrer zeitlichen Abfolge dokumentiert werden. Um den allgemeinen Anforderungen gerecht zu werden, soll im Einzelnen das Programm folgende Eigenschaften aufweisen:

- Die Benutzeroberfläche soll sich nach der Art der Arbeitsaufgabe (Konstruktion, Dokumentation, Recherche) richten.
- Die Änderungen zwischen den Versionen sollte der Bearbeiter kommentieren können und vom System optisch hervorgehoben werden. Alle Fragebögen und Einzelfragen sollen samt Änderungshistorie in einer geeigneten Datenbank gespeichert werden.
- Fragen bereits erstellter Fragebögen sollten in neue Fragebögen übernommen werden können.

¹ Insbesondere für methodische Sekundäranalysen lassen sich vorhandene Datenbestände wie die des Zentralarchivs bisher nicht nach formalen Fragekriterien durchsuchen, um z. B. Studien zu finden, bei denen eine offene Frage nach einem Gegenstand durch eine kategoriale Frage gefolgt wird. Für die Untersuchung von Kontexteffekten sind solche Recherchefragestellungen von zentraler Bedeutung. Mit einem Fragebogendokumentationsprogramm ließen sich solche Erschliessungen vorhandener Datenbestände erstmalig durchführen.

- Fragen sollen mit Daten über Fragetyp, Art der Antwortvorgaben, Fragegegenstand und Kontext der Frage (Filterführungen, Stellung im Fragebogen) gespeichert werden.
- Layout-Änderungen durch den Pretest oder Wechsel des Erhebungsmodus sollten weitgehend automatisiert werden.
- Das Programm sollte über Visualisierungen der möglichen Pfade durch den Fragebogen verfügen, um Inkonsistenzen zu vermeiden.
- Das System sollte von mehreren Benutzern gleichzeitig genutzt werden können.
- Der Fragebogen sollte in unterschiedliche Formate exportiert werden können (z. B. Word, PDF, RTF, HTML, CATI-Systeme).

3.2 Arbeitsprogramm

Da bei einer Softwareentwicklung der genaue zeitliche Ablauf der Arbeitsschritte im Voraus nur schwer festgelegt werden kann, sollen hier die grundlegenden Arbeitsschritte lediglich in der geplanten Reihenfolge dargestellt werden:

1. Entwicklung der Benutzeroberfläche
2. Bereitstellung der Fragebogendatenbank
3. Entwicklung eines Systems für gleichzeitiges Arbeiten mehrerer Benutzer
4. Implementierung eines Export- und Druckmoduls
5. Testanwendung anhand einer aktuellen Fragebogenentwicklung
6. Testanwendung anhand einiger existierender Fragebögen

Die für die Fragebogenerstellung verwendete Oberfläche richtet sich nach der benötigten Funktionalität. Die entstandenen Fragebogenversionen werden über XML-Datenbanken abgespeichert. Um das Arbeiten verschiedener Personen an diesen Datenbanken zu ermöglichen, werden WWW-Server und Java Application Server verwendet. Die Endversionen der Fragebögen werden anhand von XSL-Transformationen erzeugt.

3.2.1 Bereitstellung der Fragebogendatenbank

Um eine konsistente Datenhaltung zu ermöglichen, die sowohl mehreren Benutzern die Bearbeitung von Fragebögen ermöglicht, als auch die Datenintegrität der Daten sicherstellt, wird die Speicherung der Daten in einer Datenbank vorgeschlagen. Dabei liegt es nahe XML-Datenbanken statt relationaler als Basis einzusetzen. Über diesen Datenbanktyp gibt es jedoch heute noch relativ wenig Erfahrung; eine experimentelle Realisierung, im Rahmen eines Projektmodells, die allein tragfähige Aussagen über bestehende Vorteile bzw. Nachteile in dem konkreten Anwendungsgebiet des Projekts

erlaubt, sollte deshalb der wichtigste Projektbestandteil sein. Bei XML-Datenbanken existieren im OpenSource-Bereich (Xindice und eXist) noch keine „stabilen“ Versionen. Auch unterstützen diese Datenbanken das Update einzelner Dokumentbestandteile noch nicht sehr effizient. Kommerzielle Produkte haben hohe Lizenzgebühren, ein deutlicher Hinderungsgrund für Forschung und Lehre, wohl aber eine vernünftige Lösung für organisatorische Strukturen, die größere Mengen von Fragebogenentwicklungen zentral verwalten.

3.2.2 Implementierung eines Export- und Druckmoduls

Zur Erzeugung von Export- oder Druckversionen hat sich in anderen Kontexten bereits die Transformation von XML-Daten über XSL-Transformationen bewährt. Am IZ wird so z. B. ohne manuellen Eingriff die soFid Themendokumentation erzeugt. Dabei werden pro soFid-Themen mehrere hundert Datenbanknachweise aus der Datenbank gelesen, in XML formatiert und über XSLT nach RTF transformiert. Das Layout der RTF-Datei entspricht exakt dem früher direkt in Word erzeugten Layout, so dass zum Erzeugen der PDF-Datei für den Druck nur das Öffnen der RTF-Datei in Word und Drucken auf dem AcrobatDistiller notwendig ist. Zur Erzeugung der Fragebögen wird daher ebenfalls vorgeschlagen, die in der Datenbank gehaltenen Daten auszulesen, nach XML zu konvertieren und dann mittels XSLT in RTF oder – falls möglich – direkt nach PDF zu transformieren. Letzte Korrekturen, z. B. das Einfügen von Logos oder komplizierter Formatierungen, können dann auf der Basis der automatisch generierten RTF-Datei in Word durchgeführt werden. Bei Änderungen des Fragebogens können die einzelnen Fragen neu im RTF-Format generiert und gegebenenfalls einzeln in die bestehende Word-Datei eingefügt werden. Die Anforderungen an eine flexible, konsistente und nachvollziehbare Bearbeitung der Fragebögen würde durch das skizzierte Verfahren ohne Abstriche gewährleistet. Der endgültige Fragebogen könnte – je nach Anforderungen an das Layout – in RTF zur Nachbearbeitung in Word oder direkt zum Druck als PDF-Datei generiert werden. Daneben sind beliebige andere Ausgabeformate denkbar, z. B. CATI/CAPI-Formate oder die automatische Generierung von HTML-Fragebögen für Online-Befragungen. Die damit erfassten Daten könnten über eine zusätzliche Systemkomponente direkt zur Auswertung in eine Datenbank gespeichert werden.

3.2.3 Testanwendung anhand einer aktuellen Fragebogenentwicklung

Als Testmaterial für das Dokumentationssystem soll die bislang in QDDS vorliegende Dokumentation des Defect-Projekts verwendet werden. Sollte eine ähnlich detaillierte Instrumentenentwicklung während der Projektlaufzeit stattfinden (z. B. innerhalb des Mannheimer Mini-Panels zur Scheidungsforschung), wird die Arbeitsgruppe versuchen, die Instrumentenentwicklung durch das Programm zu unterstützen.

3.2.4 Testanwendung anhand einiger existierender Fragebögen

Als naheliegende Anwendung für existierende Fragebögen kommen die Standarddatensätze der empirischen Sozialforschung (SOEP, ALLBUS) in Frage. Die Kodierung einer Allbus- und einer SOEP-Erhebungen würde die Brauchbarkeit des Systems belegen¹.

¹ Eine Kodierung aller SOEP und ALLBUS-Erhebung würde den Rahmen eines Tests sprengen. Sollte das Projekt erfolgreich verlaufen, wird die Arbeitsgruppe an der Universität Konstanz einen entsprechenden Folgeantrag stellen. Vorgespräche mit der SOEP-Gruppe haben bereits stattgefunden.

4 Literatur

- Coutts-Heller,E. (2002): Context effects in the measurement of subjective probabilities in surveys, Diplomarbeit, Universität Konstanz.
- Kreuter,F.(2002): Kriminalitätsfurcht: Messung und methodische Probleme, Opladen: Leske + Budrich.
- Kreuter,F./Schnell,R. (2000): Die Bibliothek in den Augen der Studierenden. Ergebnisse der Befragung der Studierenden an der Universität Konstanz, <http://www.ub.uni-konstanz.de/kops/volltexte/2000/500/>
- Kreuter,F./Schnell,R./Kopp,J. (2000): Absolventenbefragung des Fachbereichs Politik- und Verwaltungswissenschaft der Universität Konstanz, <http://www.ub.uni-konstanz.de/kops/volltexte/2000/521/>
- Kreuter,F./Thume,D./Coutts,E./Venningen,H./Schnell,R. (2005): Sampling-Errors und Nonsampling-Errors in komplexen Bevölkerungsstichproben. Das „Defect“-Projekt. Ein Methodenbericht (wird 2005 im Verlag für Sozialwissenschaften erscheinen).
- Schnell,R./Kreuter,F. (2000): Das DEFECT-Projekt: Sampling Errors und Nonsampling Errors in komplexen Bevölkerungsstichproben; in: ZUMA-Nachrichten, 47, S. 89-101.
- Schnell,R./Kreuter,F. (2001): Neue Software-Werkzeuge zur Dokumentation der Fragebogenentwicklung. ZA-Information, 48: S.56-70.
- Schnell,R./Kreuter,F. (2002): New Software tools for questionnaire development documentation, Paper presented at QDET November 2002, unveröffentlichtes Konferenz-Poster.
- Waber,A. (1998): Instrumentenentwicklung zur Bürgerbefragung Konstanz: Anwendung moderner Pretestmethoden und experimentelle Überprüfung eines Fragekomplexes, Diplomarbeit Universität Konstanz.