

Entwicklung und Implementierung einer markup-orientierten Sprache zur flexiblen Erstellung und Durchführung von WWW-Umfragen

Dorthe Lübbert¹

Eines der Grundprobleme von Online-Umfragen ist, daß zur Erstellung und Durchführung der Umfragen sowohl Programmierfähigkeiten als auch methodisches Know-How vorhanden sein müssen (vgl. Batinic, Puble, Moser 1999). In der einfachsten Variante muß ein Programm entwickelt werden, das die eingehenden Antworten auf dem Webserver speichert und in der Lage ist, sie in geeigneter Form zum Export bereitzustellen. Oft wird daneben der Einsatz von Fragefilmen und die Möglichkeit, eingehende Antworten auf ihre Stichhaltigkeit zu prüfen (Plausibilitätskontrollen), gewünscht. Derartige Programme müssen mit „traditionellen“ Mitteln für jede Umfrage neu erstellt werden, da sich die Variablenstruktur von Umfrage zu Umfrage ändert; ein relativ zeitaufwendiges und unter Umständen auch teures Unterfangen, z.B. wenn Programmierer bezahlt werden müssen.

Lösungsansatz

Das Softwarepaket MOOSE² versucht, das Problem mit der folgenden Idee zu lösen: Entwickelt wurde eine Satzsprache, mit der Umfragen umfassend beschrieben werden können, und zwar in einer Form, die einerseits von Computern verstanden und andererseits von Menschen leicht erzeugt werden kann. Die Umfragen werden in dieser Satzsprache formuliert und auf den Webserver übertragen. Dort sorgt ein Programmpaket in der Programmiersprache Perl für die Durchführung und Verwaltung der Umfrage. Die einzelnen Bildschirmseiten einer Umfrage werden dynamisch aus der Satzdatei generiert, Steueranweisungen ausgewertet und die eingehenden Antworten gespeichert. Auf diese Weise kann sehr flexibel und schnell eine beliebige Anzahl verschiedener Umfragen durchgeführt werden, ohne daß Änderungen am Programmpaket vorgenommen werden müssen.

1 Dorthe Lübbert, Heinrichstr. 40, 44137 Dortmund, E-Mail: Dorthe@Luebbert.net. Dieser Beitrag ist im Rahmen einer Diplomarbeit an der Ruhr-Universität Bochum, Fakultät für Sozialwissenschaft, bei Prof. Dr. Werner Voß, entstanden.
2 MOOSE kürzt *Markup Organization Of Survey Environments*, Markup-Organisation von Umfrageumgebungen, ab.

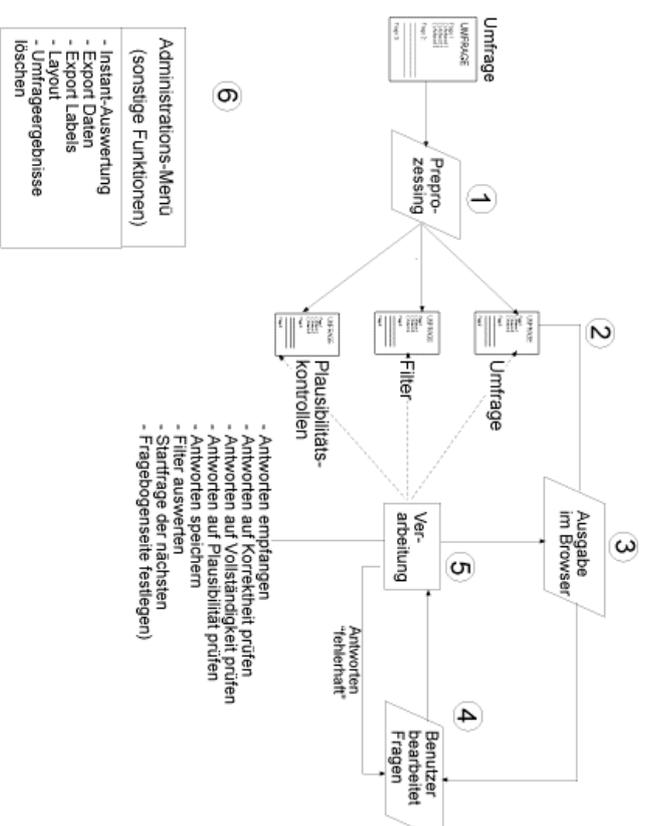


Abbildung 1: Systemarchitektur

Systemarchitektur

Abbildung 1 zeigt überblicksartig, wie die einzelnen Komponenten von MOOSE zusammenarbeiten. Die Umfrage-Satzdatei wird zunächst einem Preprozessing unterzogen, bei dem die Umfrage initialisiert und für die Verarbeitung auf dem Webserver vorbereitet wird (z.B. werden Umlaute und Sonderzeichen kodiert, in XML formulierte Fragefilter und Plausibilitätskontrollen in separate Perl-Programme umgesetzt).

Die Schritte 3 bis 6 beschreiben die Durchführung der Umfrage. Bei jedem neuen Befragungsteilnehmer wird die Satzdatei gelesen und daraus ein Webfragebogen erstellt. Der Fragebogen wird dann auf den Computer des Befragten übertragen (Schritt 4).

Nachdem der Befragungsteilnehmer den ausgefüllten Fragebogen zum Webserver zurückgeschickt hat, werden die eingegangenen Antworten in Schritt 5 geprüft und gespeichert. In der Phase „Prüfung“ wird kontrolliert, ob bei halboffenen Fragen mit Einfachauswahl nur eine Antwortalternative ausgewählt wurde, ob alle als „notwendig“ markierten Fragen beantwortet wurden und ob die Antworten in eventuell vorgegebenen Antwortbereichen liegen (Plausibilitätskontrollen). Wenn eine dieser Kontrollen fehlschlägt, wird eine entsprechende Fehlerseite erzeugt und der Befragungsteilnehmer zur Korrektur seiner Antworten aufgefordert.

Danach wird die nächste Fragebogenseite generiert, wobei für die Festlegung der „Startfrage“ die Daten aus Filterfragen berücksichtigt werden.

Ergänzt wird das Umfrageprogramm um einige weitere Funktionen, die zur eigentlichen Durchführung der Umfrage nicht zwingend erforderlich sind. Zugang zu diesen Funktionen erhält der Umfrageleiter über ein Administrationsmenü, das er über seinen Browser bedient.

Das globale Layout der Umfrage wird durch ein separates Programm festgelegt. Dazu gehört zum Beispiel die Entscheidung, welchen Copyrightvermerk die generierten Auswertungsgrafiken haben sollen, aber auch die individuelle Festlegung von Fehlermeldungen für fehlerhaft beantwortete Fragen.

Die Instant-Auswertung führt jederzeit eine deskriptive Analyse aller eingegangenen Antworten durch und stellt diese grafisch und tabellarisch dar. Berechnet werden die Häufigkeitsverteilungen aller metrischen Variablen, die grafische Darstellung erfolgt in Stabdiagrammen.

Die gesammelten Daten können mit einem Exportprogramm in anderen Anwendungen weiterverarbeitet werden. Dazu bringt das Export-Script die Datenmatrix in das gängige Austauschformat CSV. Die Datenspalten werden durch Sonderzeichen, z.B. Tabzeichen, Komma oder Semikolons getrennt. Ein zusätzliches Skript erstellt Programmcode in der Syntaxsprache von SPSS, mit dem die Variablennamen und Wertebel in SPSS eingelesen werden können.

Probeweise eingegebene Antworten können über eine weitere Funktion im Administrationsmenü gelöscht werden.

Die Satzsprache

Als Grundlage für die Satzsprache wurde XML gewählt. XML steht für *Extended Markup Language* und gehört wie HTML zu den Markup-Sprachen, die dadurch charakterisiert sind, daß sie aus Klartext mit eingeschobenen Befehlen (*Tags*) bestehen. XML bietet im Gegensatz zu HTML die Möglichkeit, eigene Befehle zu definieren und damit Texte logisch zu strukturieren. Da XML HTML in seiner Struktur sehr ähnlich ist, können XML-Texte von Benutzern mit etwas HTML-Erfahrung relativ leicht erstellt werden.

```
<simpleradio number="10" text="Wie alt sind Sie?">
<data>
<details vname="alter" value="1" label="unter 20 Jahre alt"/>
<details vname="alter" value="2" label="20 bis unter 30 Jahre alt"/>
<details vname="alter" value="3" label="30 bis unter 40 Jahre alt"/>
<details vname="alter" value="4" label="über 40 Jahre alt"/>
</data>
</simpleradio>
```

Abbildung 2: Beispiel für die "Programmierung" eine Frage mit Einfachauswahl in XML

Abbildung 2 zeigt exemplarisch den XML-Code einer Frage mit Einfachauswahl. `simpleradio` ist das Schlüsselwort für eine Frage mit Einfachauswahl³. Abbildung 3 stellt dar, wie die Frage beim Befragen angezeigt wird.

Als weiteres Attribut wird die fortlaufende Nummer dieser Frage festgelegt (`number="10"`), „Wie alt sind Sie?“ ist also die zehnte Frage der Umfrage. Fragen können auf mehrere Webseiten verteilt werden, einzelne Fragen oder ganze Blöcke können durch Filter übersprungen werden. An Hand der Nummer wird entschieden, ob die Frage im „aktiven Teil“ der Umfrage steht und ausgegeben werden muß.

Abbildung 3: Darstellung des XML-Codes im Browser

Wie alt sind Sie?

unter 20 Jahre

zwischen 20 und 30 Jahre

zwischen 30 und 40 Jahre

zwischen 40 und 50 Jahre

über 50 Jahre

Zwischen `<data>` und `</data>` werden die möglichen Antworten auf die Frage festgelegt. Die Antwortkategorien werden einzeln innerhalb von `<details>`-Anweisungen definiert. `<details vname="alter" value="1" label="unter 20 Jahre alt">` legt die erste Antwortalternative fest. Die gegebene Antwort wird in der Variablen `alter` gespeichert, die die Ausprägung 1 annimmt, wenn die erste Antwortalternative ausgewählt wurde.

Ausblick

MOOSE stellt in seiner jetzigen Form komplexe Funktionen zur Durchführung von WWW-Umfragen zur Verfügung, allerdings muß die Satzdatei manuell mit einem Texteditor erstellt werden. Diese Vorgehensweise erschwert den Zugang für Umfrageleiter, die nur über geringe technische Kompetenz verfügen und bsp. keinerlei Erfahrung mit HTML mitbringen. Wünschenswert wäre es deshalb, MOOSE mit einem Umfrageeditor zu koppeln, über den die Umfrage interaktiv „zusammengeklickt“ werden kann.

Da MOOSE komplett im Quellcode vorliegt, können Änderungen und Ergänzungen relativ leicht vorgenommen werden. Wollte man z.B. automatisch Java-Script-Code zu einem bestimmten Fragetyp ausgeben, müßte man lediglich einige Print-Anweisungen in das Programm einfügen.

³ Zur Bezeichnung `simpleradio`: `simple` steht für für simple choice (Einfachauswahl) und `radio` für Radiobuttons.

Literatur

- Batinić, B.; Puhle, B.; Moser, K. (1999): *Der WWW-Fragebogen-Generator (WFG)*.
In: Batinić, B.; Werner, A.; Gräf, L.; Bandilla, W. (Hrsg.): *Online Research. Methoden, Anwendungen und Ergebnisse*. Göttingen.
- Dillman, D.; Tortora, R.; Bowker, D. (1998): *Principles for Constructing Web Surveys*. [WWW document]
URL: <http://survey.sestrc.wsu.edu/dillman/papers/websurveypr.pdf>.
- Gadeib, A. (1999): *Ansprüche und Entwicklung eines Systems zur Befragung über das World Wide Web*. In Batinić, B.; Werner, A.; Gräf, L.; Bandilla, W. (Hrsg.): *Online Research. Methoden, Anwendungen und Ergebnisse*. Göttingen.
- Leventhal, M., Lewis, D., Fuchs, M. (1998): *Designing XML Internet Applications*, Upper saddle River, NJ.