Hypermedia als Visualisierungs- und Suchoberfläche eines Umweltinformationssystems

W. Sinowski, M. Knorrenschild, E. Ernst, A. Fieger[1]

1 Einleitung

Große Netze, wie z.B. das Internet, eignen sich als Kommunikationsebene in besonderem Maß für Informationssysteme, die einen großen Nutzerkreis ansprechen wollen. Die Vielzahl an verschiedenen Hard- und Softwaresystemen führt jedoch zu Problemen bei der Gestaltung einer Visualisierungsoberfläche für die Informationen. Übertragungsprotokolle wie Telnet, die nahezu jedem für den Online-Informationsaustausch im Internet zur Verfügung stehen, sind in ihren Möglichkeiten sehr begrenzt. Dies verursachte in der Regel benutzerunfreundliche Oberflächen, deren Handhabung sehr gewöhnungsbedürftig ist. Seit wenigen Jahren eröffnet in diesem Bereich World Wide Web über HTML völlig neue Möglichkeiten. Für die Visualisierung der Informationen aus einem Umweltinformationssystem ergeben sich dadurch neue Aspekte, die bedacht und diskutiert werden sollten.

2 Verbindung von Datenbank und HTML-Seiten

HTML-Seiten sind zwar zur Visualisierung von Informationen sehr geeignet. Sie ersetzen jedoch nicht die Leistungsfähigkeit einer Datenbank zur Verwaltung der Informationen. Es liegt daher nahe, Lösungen in einer Verknüpfung der Datenbank zur Informationsverwaltung mit HTML-Seiten als Visualisierungsoberfläche zu suchen (Rowe, 1995). Für die Generierung der HTML-Seiten können dabei unterschiedliche Wege verfolgt werden. Die HTML-Seiten können:

- einmal erstellt und ab diesem Zeitpunkt unverändert bleiben (statisch).
- in bestimmten Zeitintervallen mit den Inhalten der Datenbank aktualisiert werden (intervalldynamisch).
- zum Zeitpunkt eines Nutzerzugriffs, abhängig von den Eingaben des Nutzers, generiert werden (dynamisch).

Neben den eventuell unterschiedlichen Erfordernissen an die Reaktionszeit für eine Aktualisierung hängen die Vor- und Nachteile dieser drei Generierungswege darüber hinaus von den Funktionen der einzelnen HTML-Seiten ab.
Bei einem Umweltinformationssystem sind diese Funktionen vor allem die

- Darstellung von Suchwegen zur Information,
- Darstellung der Information an sich,

---

[1] GSF - Forschungszentrum für Umwelt und Gesundheit, Projektgruppe Umweltgefährungspotentiale von Chemikalien, Neuherberg, 85764 Oberschleißheim

- Darstellung von Informationen zum System selbst oder zur
- Eingabe neuer Informationen an das System.

3 Charakterisierung von UFIS

UFIS stellt Informationen zur ökologischen Forschung über Modelle und modellrelevante Daten bereit und ist seit Februar 1995 über das WWW öffentlich zugänglich (http://www.gsf.de/UFIS/ufis).
Die von UFIS angesprochenen Nutzergruppen sind:

- Fördernde Institutionen der ökologischen Forschung,
- Wissenschaftler, die Modelle entwickeln, nutzen und verbinden wollen, und komplexe Ursache-Wirkungs-Beziehungen in unserer Umwelt simulieren, prognostizieren und bewerten wollen.

Diese Nutzergruppen kommen dabei aus verschiedenen Fachgebieten und Interessenskreisen und benötigen daher die Informationen aus unterschiedlichen Sichtweisen und mit unterschiedlicher Detailtiefe.
Das Informationssystem teilt sich in Informationen zu Modellen und Informationen zu Daten, wobei zur Zeit der Teil über die Informationen zu Modellen wesentlich weiter fortgeschritten ist (Lenz u.a. 1995). Ziel ist es, mit den in Form von Metainformationen gesammelten Modell- und Datendokumentationen ein breites Spektrum der Ökologie, das sich über alle räumlichen, zeitlichen und ökosystemaren Skalierungsebenen erstreckt, abzudecken. Dabei wird eine Dokumentationstiefe angestrebt, die detailliert genug ist, Abfragen über Modellverknüpfung oder über den Abgleich zwischen dem Datenbedarf von Modellen und dem Datenangebot aus Meßdatenbanken zu ermöglichen. Aufgrund dieser Dokumentationstiefe erfährt UFIS nur einen relativ langsamen Input, der nur einen wöchentlichen Aktualisierungszyklus notwendig macht.
Die Objekte der Modelldokumentationen, die über WWW visualisiert werden, sind:

- der allgemeine Teil der Modellbeschreibung
- die Informationen zu den Modellgleichungen
- die Informationen zu den Modellparametern und Input- und Outputgrößen.

Zusätzlich werden allgemeinen Projektinformationen auf WWW angeboten.

4 Visualisierung der Informationsobjekte

4.1 Visualisierung der Modelldokumentationen

Die Dokumentationsinhalte zu den oben beschriebenen Objekten werden aus der UFIS-Datenbank unter INGRES über ein C++-Programm über embedded SQL ausgelesen und in strukturierte HTML-Texte übersetzt. Jede Modell-, jede Gleichungs- und Größendokumentation wird einzeln als HTML-Dokument abgelegt. Der Dokumentname enthält den eindeutigen ID des Objektes, unter dem es auch in der Datenbank verwaltet wird. Die Informationen zu einem Modell sind dabei in

so die Modellgleichungen grafisch in ihren mathematischen Formulierungen sehen. Die zugrundeliegende LaTeX - Notation bleibt auskommentiert im HTML-Quelltext verfügbar.

4.3 Visualisierung der Informationen über UFIS selbst

Informationen über das UFIS-Projekt selbst sind in jedem Fall unabhängig von den wachsenden und sich verändernden Inhalten der Datenbank. Sie werden daher nur auf rein statischen HTML-Seiten vorgehalten.

5 Darstellung von Suchwegen zu den Informationsobjekten

Da die Nutzer von UFIS aus verschiedenen Fachgebieten und Interessenskreisen kommen, bietet UFIS mehrere Suchapplikationen die Informationsobjekte zu finden an, damit jeder Nutzer seinen individuellen Suchweg verfolgen kann:

- Namensliste der Modelle.
  Kennt der Nutzer konkret den Modellnamen, für das er weitere Informationen benötigt, so bietet eine einfache alphabetische sortierte Namensliste der in UFIS vorhandenen Modelle einen schnellen Zugriff auf die entsprechende Dokumentation. Die Modellnamensliste ist im momentanen Stand von UFIS mit 37 Modellen noch handhabbar. Mit zunehmender Anzahl an Modellen wird jedoch diese Namensliste durch ein Suchformular ersetzt werden müssen.
- Stichwortlisten nach den Themen Modellapplikation, Modelltyp, Kurzbeschreibungen, Gleichungen und Größen.
  Wenn dem Nutzer bekannt ist, nach welchem Aspekt der Modelldokumentation und zu welchem Fachgebiet er Informationen sucht, so bieten diese alphabetisch sortierten Stichwortlisten Hilfe. Über die Auswahl eines Stichwortes wird eine Liste der dazu passenden Informationsobjekte aufgerufen. Ein Nachteil dieses Suchweges ist, daß die Stichworte unstrukturiert und frei formuliert sind, d.h. in der Regel aus mehreren Begriffen zusammengesetzt sind. Dies erschwert das Finden eines für die jeweilige Suchanfrage passenden Stichwortes.
- Suchbaum von Schlüsselwörtern, der bis zu den Modellen, Gleichungen und Größen führt.
  Der Suchbaum über Schlüsselwörter ermöglicht im Gegensatz zu der unstrukturierten Stichwortliste eine hierarchisch aufgebaute Suche nach bestimmten fachlichen Themengebieten und Aspekten. Die Schlüsselwörter werden aus den frei vergebenen Stichwörtern, die aus mehreren Begriffen zusammengesetzt sein können, herausgesucht (Abb. 2). Entsprechend diesen Schlüsselwörtern werden die Stichwörter den jeweiligen Positionen im hierarchischen Schlüsselwortbaum zugeordnet. Die Möglichkeit von Querverweisen in WWW wird hier dazu genutzt, jederzeit von einem Themenpfad im Schlüsselwortbaum zu einem anderen im Stichwort enthaltenen Schlüsselwort zu springen. Wie bei den

Stichwortlisten führt schließlich ein gefundenes Stichwort wieder zu einer Liste von dazu passenden Modellen, Gleichungen und Größen. Über die zu erwartende Trefferanzahl wird jedoch schon vorab informiert, um einer Demotivation der Nutzer durch Suchergebnisse ohne echten Inhalt vorzubeugen.

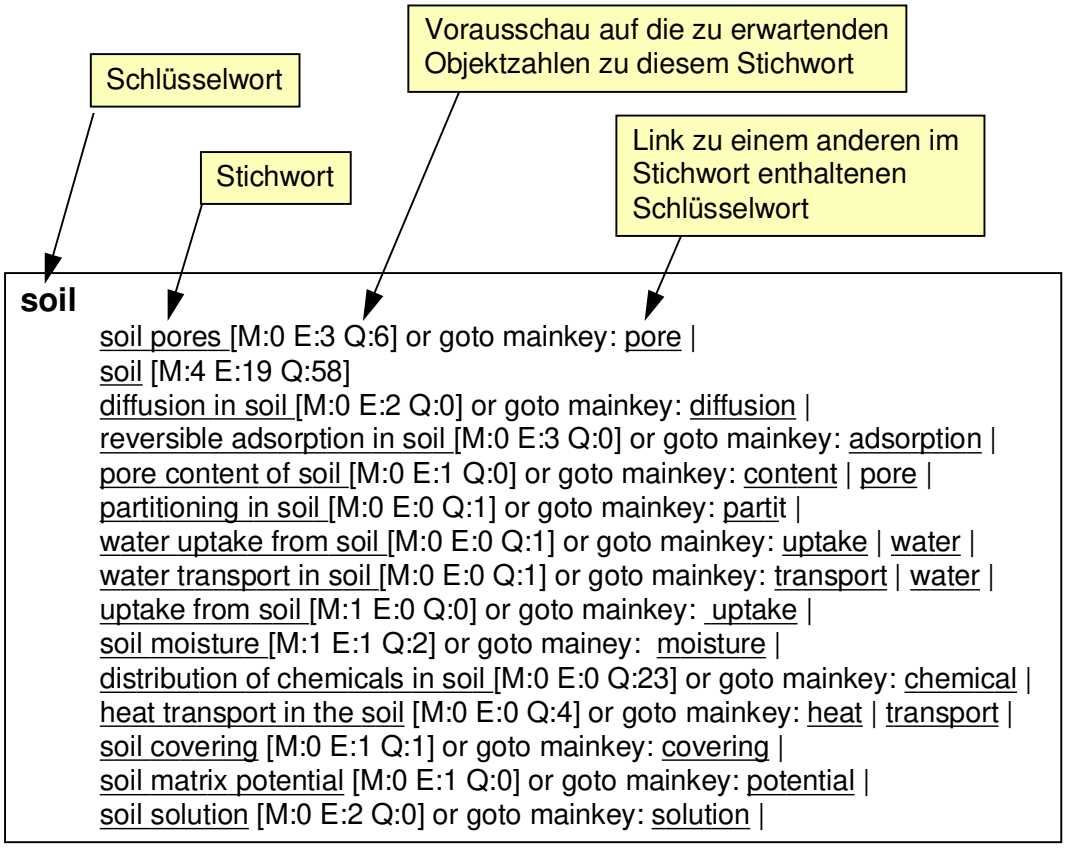

Abb. 2: Beispiel aus dem Stichwort-Suchbaum zu Modellen, Gleichungen und Größen.

- systemökologischer Thesaurus.
  Bei der Suche nach Informationsobjekten der ökologischen Forschung interessiert oft der Aspekt, auf welche Standortgegebenheiten und/oder Organisationsniveau die Aussage des Informationsobjektes (bei UFIS z.B. des Modelles) gültig ist. Dazu wurde auf der Grundlage von Miller (1975) und Springstubbe (1993) ein systemökologischer Thesaurus in das Informationssystem integriert. Dieser Thesaurus wurde mit seiner hierarchischen Struktur auch auf WWW abgebildet und kann für die Suche nach Modellen verwendet werden. Die Modelle sind dabei sowohl auf der untersten noch gültigen Organisationsebene aufgelistet, als auch auf den darüber liegenden Hierarchieästen.

Die Struktur der Suchwege ist dabei weitgehend statisch. Die Verweise zum entsprechenden Suchergebnis sowie die Inhalte des Suchergebnisses selbst werden dagegen in regelmäßigen Abständen über Programme im Hintergrund aktualisiert.
Gegenüber einer voll dynamischen Suchanfrage über ein Anfrageformular, wie es einige vergleichbare Informationssysteme auf dem WWW anbieten (z.B.: CAMASE: http://www.bib.wau.nl:80/camase/srch-cms.html, ECOBAS http://-dino.wiz.uni-kassel.de/ecobas.html), bietet dies für den Anwender folgende Vorteile:

- Es kann je nach Denkweise und Fachrichtung des Anwenders der von ihm bevorzugte Suchweg verfolgt werden.
- Eine ausgewogenere Anzahl von Suchergebnissen wird erreicht. Bei dynamischen Suchseiten, in denen über die Eingabe von Stichwörtern die entsprechenden Inhalte ausgegeben werden, führt dies bei Umweltinformationssystemen mit breitgefächertem Angebot an Themen oft zu einer sehr unausgewogenen Anzahl von Suchergebnissen. Ein zu großes Angebot an möglichen Stichwörtern führt zu Unsicherheiten, welches Stichwort am treffendsten die gewünschte Information bringt. Ein zu kleines Angebot an Stichwörtern dagegen führt in der Regel zu einem zu hohen Angebot an möglichen Informationen.
- Zudem müssen bei den dynamischen Suchseiten die möglichen Stichwörter und Abfragekombinationen dem Anwender bekannt sein und von ihm beherrscht werden. Eine vorherige Kenntnis der möglichen Stichwörter, die zu einer erfolgreichen Suchanfrage führen, ist dagegen bei einer Konzeption, wie sie am Beispiel UFIS aufgezeigt wurde, nicht nötig.
- Schnellere Zugriffe zur Information, da Suchanfragen nicht über eine Datenbankabfrage laufen müssen.

6 Eingabe neuer Metainformation in UFIS über WWW

Seit kurzem bietet UFIS auch die Möglichkeit einer Eingabe der Modell- und Datendokumentationen über HTML-Formulare in WWW an. Dies hat den Vorteil, daß die Anwender eine Eingabeoberfläche erhalten, die über den HTML-Browser an ihre gewohnte Oberfläche angepaßt wird. Auf der Seite des Anbieters (d.h. UFIS) ergaben sich daraus folgende Problemstellungen:

- Die Fragebögen sind relativ umfangreich, so daß das Ausfüllen und Abschicken in einer WWW-Sitzung nicht zumutbar ist. Daher müssen die Inhalte zwischenspeicherbar und wieder abrufbar sein. Dies ist im Konzept der HTML-Formulare nicht ohne weiteres vorgesehen.
- Abhängig vom Modell oder Datensatz wiederholen sich bestimmte Objekte mehrmals. Wie oft sich diese Objekte wiederholen, ergibt sich jedoch erst im Laufe des Ausfüllens. Dies macht die Gestaltung eines ganzen Fragebogensystems aus mehreren, aufeinander verweisenden HTML-Seiten notwendig.
- Das Einfügen von Begriffen aus hierarchisch verzweigten Thesauri ist automatisiert über HTML-Formulare schwer realisierbar.

- Die bisher benutzten INGRES-Applikation für die Eingabe kann nicht mehr verwendet werden. Neue Einlese- und Plausibilisierungsroutinen, die die als ASCII-Dateien abgelegten Inhalte in die INGRES-Datenbank überführen, sind erforderlich.

Bis auf das Einfügen von Begriffen aus hierarchisch verzweigten Thesauri konnten die oben aufgeführten Problempunkte bereits über Perl-Programme bewältigt werden und damit neben der Eingabeapplikation über INGRES eine benutzerfreundlichere Alternative angeboten werden.

7 Schlußfolgerungen

- WWW hat sich in den letzten Jahren zu einem wichtigen Visualisierungsinstrument für Informationen im Internet entwickelt. Es eignet sich daher in hohem Maße als Visualisierungsoberfläche von frei zugänglichen Datenbankinhalten, da dadurch ohne großen Entwicklungsaufwand eine weitgehend systemunabhängige und benutzerfreundliche Visualisierungsoberfläche für die Datenbankinhalte gegeben ist.
- Intervalldynamische und dynamische HTML-Seiten sind in ihrer Realisierung zunächst aufwendiger als statische Seiten, da hinter ihnen ein Programm steht, das erstellt und gepflegt werden muß. Auf veränderte Inhalte kann jedoch flexibler reagiert werden.
- Dynamische Seiten haben gegenüber intervalldynamischen Seiten den Vorteil einer Online-Aktualisierung. Dynamische Suchseiten mit einzugebenden Schlüsselwörtern haben jedoch den Nachteil der Unübersichtlichkeit bei zu großen Eingabemöglichkeiten und von oft unausgewogenen Suchergebnissen.
- Statische HTML-Seiten lassen sich mit wenig Aufwand gestalten. Sie eignen sich vor allem für Information, die kaum einer Aktualisierung bedürfen.
- UFIS ist ein anschauliches Beispiel, wie je nach Funktion einer HTML-Seite unterschiedliche Arten der HTML-Seitengenerierung zum Einsatz kommen.

8 Literatur

Lenz R./ Knorrenschild M./ Herderich C./ Forster E./ Springstubbe O./ Benz J./ Asshoff M./ Windhorst W. (1995): An information system on ecological models. Technical Report 27/94, GSF Forschungszentrum für Umwelt und Gesundheit, Neuherberg.

Lenz, R./ Haber, W. (1996): Classification theory of ecological systems - are the generalizations only the exceptions?, in: Bulletin of the Ecological Society of America, 77, S. 62-64.

Miller G. T. (1975): Living in the Environment - Concepts, Problems and Alternatives. Wadsworth Publ.

Rowe J. (1995): Building Internet Database Servers with CGI. New Riders Publishing (http://www.ncp.com/nrp), ISBN 1-56205-573-9

Springstubbe O. (1993): Grundzüge der Klassifikation und systematischen Beschreibung ökologischer Systeme als Gliederungsprinzip für ein Informationssystem ökologisch relevanter Daten und Modelle. diploma thesis, Techn. Univ. Munich, 1993.

Das diesem Beitrag zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministers für Forschung und Technologie unter dem Förderkennzeichen 0339349A gefördert.