

## 4 Benutzermodellierung in adaptiven Web-Anwendungen

*Hannes Meyer*

### 4.1 Einführung und Begriffsdefinitionen

Mit wachsender Informationsflut im World Wide Web wird die Filterung interessanter und wichtiger Informationen für die Benutzer zunehmend schwieriger. Personalisierte Softwaresysteme können dabei helfen, dem Benutzer den Zugang zu individuell relevanten Informationen zu erleichtern und die Informationen auf eine angemessene Weise zu präsentieren. Dazu müssen diese Systeme in der Lage sein, den Benutzer kennen zu lernen und Daten über ihn zu sammeln. Dieser Prozess wird als Benutzermodellierung bezeichnet. In diesem Kapitel werden unterschiedliche Aspekte der Benutzermodellierung behandelt. Zunächst wird auf Techniken der Datengewinnung eingegangen, danach mögliche Inhalte von Benutzermodellen beschrieben und schließlich unterschiedliche Möglichkeiten der formalen Repräsentation von Benutzermodellen besprochen.

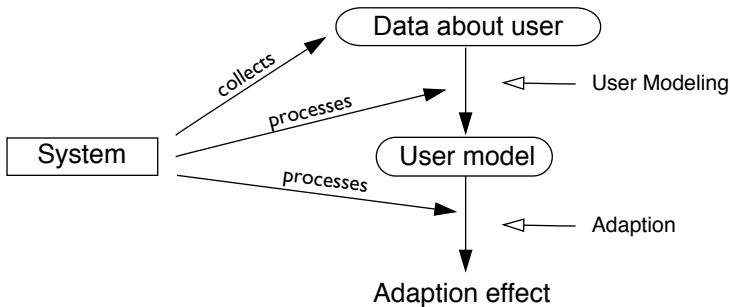
**Adaptive Systeme.** Zunächst wird hier eine Abgrenzung der Begriffe adaptiv und adaptierbar durchgeführt. Adaptierbare Softwaresysteme können durch den Benutzer selbst aktiv an seine Bedürfnisse angepasst werden (Kobsa, 2004). Auch Websites lassen sich derzeit bereits an persönliche Anforderungen anpassen, beispielsweise bei individuellen Startseiten wie iGoogle<sup>1</sup> oder Pageflakes<sup>2</sup>. Diese Art der Personalisierung erfordert vom Benutzer einen gewissen Aufwand und lässt sich daher nur begrenzt in größerem Maßstab

---

<sup>1</sup><http://www.google.de/ig>, zugegriffen am 01.02.2010

<sup>2</sup><http://www.pageflakes.com>, zugegriffen am 01.02.2010

einsetzen. Im Gegensatz hierzu sind adaptive Systeme darauf ausgelegt, weitestgehend automatisiert und ohne umfangreiche Eingriffe des Benutzers einen Personalisierungseffekt zu erreichen. Personalisierte Inhalte und Präsentationsformen haben im World Wide Web (WWW) in den vergangenen Jahren deutlich an Bedeutung gewonnen und wurden besonders im E-Commerce-Bereich durch Amazon<sup>3</sup> erstmals im großem Umfang eingesetzt. Eine Bedingung der Adaptivität einer Software ist die Möglichkeit, Daten und Informationen über den Benutzer zu sammeln und zur individuellen Anpassung zu verwenden. Diese Daten werden in einem Benutzermodell festgehalten und gepflegt. In Abbildung 4.1 wird der Zusammenhang zwischen Benutzermodellierung und Adaptierung verdeutlicht.



**Abbildung 4.1:** Usermodeling-Adaption-Loop nach Brusilovsky (1996)

**Benutzermodellierung.** Um in einem interaktiven Softwaresystem einen Adaptierungseffekt auf die Bedürfnisse des Benutzers zu erreichen, müssen gewisse Daten über den Benutzer gesammelt werden. Bei den Daten kann zwischen Benutzerdaten, Benutzungsdaten und Umgebungsdaten unterschieden werden (Kobsa, 2004). Diese Daten werden in einem Benutzermodell (*User Model*) gespeichert und verwaltet, auf dessen Basis die Software Anpassungen

<sup>3</sup><http://www.amazon.com>, zugegriffen am 01.02.2010

durchführen kann. Der Aufbau eines Benutzermodells während der Interaktion mit einem Softwaresystem wird als Benutzermodellierung bezeichnet (Kobsa, 1993) und durch die Benutzermodellierungskomponente (*User Modeling Component*) der Software durchgeführt (Wahlster und Kobsa, 1989). Beim Erwerb der Benutzermodellierungsdaten wird im Allgemeinen zwischen expliziten und impliziten Techniken unterschieden (Gauch et al., 2007). Explizite Techniken verlangen einen aktiven Beitrag des Benutzers, um sein Modell zu generieren. Dies kann systemgesteuert, durch Nachfragen des Softwaresystems während oder vor der Benutzung, oder benutzergesteuert, durch Ausfüllen von Profildateien durch den Benutzer, erfolgen (Kobsa, 1993). Bei impliziten Techniken wird das Benutzermodell ohne Eingriff des Benutzers durch Beobachtung seines Nutzungsverhaltens erstellt. Im Abschnitt 4.2 werden die unterschiedlichen Techniken mit Beispielen näher erläutert. Die Repräsentation der Benutzermodelle kann sowohl auf Basis individueller Profile als auch mittels Stereotypmodellen (Rich, 1979) durchgeführt werden. Bei letzteren werden Gruppen von Benutzern basierend auf gemeinsamen Attributen zusammengefasst. Besonders bei homogenen Anwenderkreisen mit spezifischen Anwendungsbereichen kann eine stereotypisierte Modellierung zum Einsatz kommen (Kobsa, 1985). Da in Web-Anwendungen jedoch in der Regel von inhomogenen Benutzern ausgegangen werden muss, ist eine individuelle Modellierung notwendig. Eine Kombination aus beiden Varianten bietet sich an, um das Problem der anfänglichen Datenknappheit zu lösen. In Abschnitt 4.4 werden die Repräsentationsvarianten detaillierter beschrieben. Zur Verwaltung der Benutzermodelle werden spezielle Benutzermodellierungsserver verwendet.

**Einsatzbereiche.** Adaptive Hypermedia-Systeme bezeichnen Hypermedia-basierte Systeme, die mittels Benutzermodellen eine Adaption auf die Bedürfnisse des Benutzers durchführen (Brusilovsky, 1996). Adaptive Systeme existieren in vielen Bereichen wie zum Beispiel als E-Learning-Anwendungen, Online-Informationssysteme wie

E-Commerce-Anwendungen oder im Information Retrieval (Brusilovsky, 2001), zum Beispiel bei Suchmaschinen (Shen et al., 2005; Sieg et al., 2007). Die Adaption kann sich dabei auf die Inhalte der Systeme (*Adaptive Presentation*) und auf die Navigation über die Inhalte (*Adaptive Navigation Support*) erstrecken (Brusilovsky, 2001). Letzteres wird in Kapitel 5 dieses Buches ausführlich behandelt.

## 4.2 Datengewinnung zur Benutzermodellierung

Bei der Datengewinnung zum Aufbau von Benutzermodellen kann grundsätzlich zwischen expliziten und impliziten Techniken unterschieden werden. Abhängig von der Art der zu sammelnden Informationen, der Partizipations- und Auskunftsbereitschaft des Benutzers und dem angestrebten Ziel und Umfang der Adaptierung hat jede Methode Vor- und Nachteile. Daher werden in der Regel beide Bereiche in Kombination eingesetzt.

**Explizite Datengewinnung in Web-Anwendungen.** Benutzermodellierung durch explizite Angaben des Benutzers geschieht im Web in der Regel über Formulare, die ein Benutzer ausfüllen muss. Dabei können demographische Daten, Interessen oder Meinungen abgefragt werden. Insbesondere harte Fakten wie Beruf, Geburtsdatum etc. können nur auf explizite Weise gewonnen werden. Problematisch bei dieser Weise der Benutzermodellierung sind der Zeitaufwand seitens des Benutzers, die Notwendigkeit der manuellen Profilpflege, sowie fehler- oder lückenhafte Angaben des Benutzers. Aus Gründen des Datenschutzes sind viele Benutzer unter Umständen zurückhaltend beim Ausfüllen von Formularen. Viele Web-Anwendungen verwenden explizite Benutzerangaben, um die Präsentation der Inhalte zu personalisieren und interessante Inhalte vorzuschlagen (Gauch et al., 2007). Um Inhalte vorschlagen zu können, bewerten Benutzer meist ihnen bekannte Objekte, wie zum Beispiel Filme, auf einer linearen Skala. Anhand dieser Bewertungen werden mit Methoden des kollaborativen oder inhaltsbasierten

Filterns neue Objekte vorgeschlagen (Schafer et al., 2007; Pazzani und Billsus, 2007; Sarwar et al., 2001).

**Implizite Datengewinnung in Web-Anwendungen.** Implizite Datengewinnung geschieht für den Benutzer unsichtbar im Hintergrund während der Benutzung eines Systems. Zur Benutzermodellierung in Web-Anwendungen können unterschiedliche Techniken differenziert werden, Tabelle 4.1 gibt einen knappen Überblick über diese. Durch eine Überwachung des Browser-Cache können sämtliche Browsing-Aktivitäten des Benutzers in das Benutzermodell einfließen. Hierzu müssen allerdings vom Benutzer oder einer Software die Cache-Dateien hochgeladen werden. Eine alternative Vorgehensweise um den gleichen Effekt zu erzielen stellt die Einrichtung eines Proxy-Servers dar, welcher sämtliche Aktivitäten des Benutzers protokolliert und in das Modell einfließen lässt. Vorteile dieser beiden Techniken sind eine umfangreiche Datenbasis, die Möglichkeit, mehrere Websites personalisieren zu können sowie eine unkomplizierte Handhabung. Zwei weitere Gruppen stellen Software-Agenten dar, die entweder nur im Browser arbeiten oder als Desktop-Anwendung Zugriff auf den gesamten Rechner des Benutzers ermöglichen. Diese Agenten können die Arbeit des Nutzers in Echtzeit protokollieren und in ein Benutzermodell übermitteln. Da es sich um clientseitige Anwendungen handelt, lassen sich viele Desktop-Anwendungen auf Basis des Benutzermodells personalisieren. Per Übermittlung der Daten an einen Webserver können diese Daten auch zur Personalisierung von Websites verwendet werden. Desktop-Agenten haben zusätzlich den Vorteil, dass auch Dokumente, mit denen der Benutzer lokal arbeitet, analysiert werden können und sich dadurch Daten für das Benutzermodell generiert lassen. Problematisch bei Agenten ist, dass Software installiert werden muss und diese vom Benutzer verwendet werden muss. Diese Problematik stellt sich nicht bei reinen serverseitigen Ansätzen wie dem Web-Log-Mining und der Analyse von Suchmaschinen-Logs, die einen Teilbereich des Web-Usage-Mining (Mobasher, 2006) definieren. Hierbei werden Daten, die bei

der Benutzung einer Website auf dem Webserver anfallen, zum Aufbau eines Benutzermodells verwendet. Allerdings muss hierbei eine Möglichkeit zur Identifikation individueller Nutzer implementiert werden. Mögliche Techniken stellen Cookies oder Log-In-Verfahren dar. Im folgenden Abschnitt wird auf das Web-Usage-Mining detaillierter eingegangen.

Technik	Information	Reichweite	Pro	Contra
Browser Cache	Browse-Verlauf	jede Website	Benutzer muss nichts installieren	Benutzer muss seinen Browser-Cache hochladen
Proxy-Server	alle Browsing-Aktivitäten	jede Website	Benutzer kann jeden Browser verwenden	Benutzer muss einen Proxy konfigurieren
Browser-Agenten	alle Browsing-Aktivitäten	jede personalisierte Anwendung	Kann jegliche Browsing-Aktivität loggen	Software muss installiert und benutzt werden
Desktop-Agenten	alle Aktivitäten des Nutzers	jede Anwendung	Alle Dateien und Aktivitäten werden geloggt	Software muss installiert werden
Web-Log-Mining	Browsing-Aktivitäten	eine Website	Informationen über alle Nutzer werden gesammelt	Nur auf eine Website beschränkt
Suchmaschinen-Logs	Web-Suche	Seite der Suchmaschine	Sammlung und Verwendung auf einer Seite	Cookies bzw. Login erforderlich, wenig Information

**Tabelle 4.1:** Techniken impliziter Benutzermodellierung nach Gauch et al. (2007)

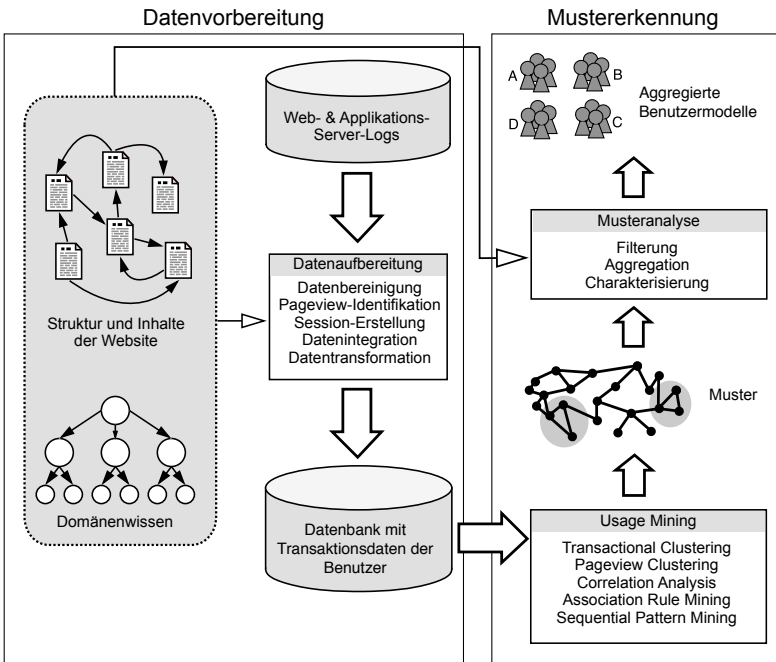


Abbildung 4.2: Web-Usage-Mining-Prozess nach Mobasher (2006)

### 4.2.1 Web-Usage-Mining

Der Prozess des Web-Usage-Mining (siehe Abbildung 4.2) gliedert sich in relativ unabhängige Phasen der Datengewinnung und -aufbereitung, Musterentdeckung (*Usage Mining*) und Musteranalyse, analog zur Vorgehensweise bei klassischem Data-Mining. Das Ziel des Web-Usage-Mining sind aggregierte Benutzermodelle, die sowohl eine Adaption der Website als auch Analysen über die Zielgruppen und deren Verhalten ermöglichen. Da es sich hierbei um ein sehr umfangreiches Thema handelt, welches nur in gewissen Bereichen

1	68.249.71.241 - - [21/Jul/2009:05:36:39 +0200] GET /Company/Hessen/Rheingau-Taunus-Kreis HTTP/1.1 200 4813 - Opera/9.63 (Windows NT 5.1; U; de) Presto/2.1.1
2	52.180.97.162 - - [21/Jul/2009:13:53:22 +0200] GET /photos/thumb/96.jpg HTTP/1.1 304 0 http://www.website.de/galleries Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_5_7; de-de) AppleWebKit/530.19.2 (KHTML, like Gecko)
3	52.180.97.162 - - [21/Jul/2009:14:28:17 +0200] GET /suche?search_q=79395&search%5Bradius%5D=5 HTTP/1.1 301 167 http://www.website.de/ Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_5_7; de-de) AppleWebKit/530.19.2 (KHTML, like Gecko)
4	52.180.97.162 - - [21/Jul/2009:14:46:08 +0200] GET /suche/5/km_um/Gechingen/16870 HTTP/1.1 200 8007 http://website.de/galleries Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_5_7; de-de) AppleWebKit/530.19.2 (KHTML, like Gecko)

Tabelle 4.2: Beispiel eines Server-Log

Überschneidungen mit dem Thema Benutzermodellierung aufweist, wird an dieser Stelle lediglich die erste Phase der Datengewinnung und Datenaufbereitung näher beschrieben. Detailliertere Ausführungen zum Thema finden sich bei Mobasher (2007, 2006, 2004); Omari und Conrad (2006); Pierrakos et al. (2003).

In der ersten Phase der Datengewinnung werden Daten aus unterschiedlichen Quellen gesammelt. Die primäre Quelle der *Benutzungsdaten* stellen Web-Server-Logs dar (siehe Tabelle 4.2), in denen ein Webserver jeden Aufruf einer Webseite oder eines Dokuments protokolliert. Anhand der IP-Adresse und Daten des User-Agent lassen sich individuelle Benutzer identifizieren, deren Klickverhalten auf der Website durch Analyse der aufgerufenen URLs und deren *Referrer* bestimmen. Einzelne Aufrufe von Seiten werden als *Pageviews* bezeichnet, eine Sammlung von Pageviews während eines Besuches einer Website stellt eine *Session* dar. Zusammengeführt werden diese Daten mit Informationen über den Inhalt der einzelnen Seiten, inklusive eingebundener Grafiken, Multimediainhalten, Meta-Daten und

Hierarchieebenen sowie über den strukturellen Aufbau der Website. Zu letzterem zählen Verlinkungen zwischen verschiedenen Seiten der Website, wie sie in Sitemaps dargestellt werden. Bei dynamischen Websites, die über ein Log-In-System verfügen, können unter Umständen auch Daten des Benutzerprofils wie Bewertungen, demographische Daten sowie weitere explizit oder implizit gewonnene Erkenntnisse über den Benutzer mit einfließen. Anschließend müssen die Daten bereinigt werden, da Logfiles auch Informationen über Besuche von Suchmaschinen-Bots beinhalten, eingebettete Inhalte wie Grafiken, Scripts oder Stylesheets nicht in die Analyse mit einfließen sollen und gewisse Inhalte der Logfiles nicht interessant sind. Zudem können die Benutzungspfade der Nutzer lückenhaft sein, da über clientseitiges Caching nicht jede Seite vom Server abgerufen werden muss und Benutzer unerkannt in ihrem Browser-Verlauf zurück navigieren können. Anhand der Logfiles und Kenntnissen über den Aufbau der Website können diese fehlenden Informationen jedoch rekonstruiert werden.

Anhand der Benutzerdaten der resultierenden Transaktionsdatenbank werden mittels Data-Mining-Techniken Analysen durchgeführt, um signifikante Muster in der Benutzung der Website zu erkennen. Durch die Zusammenführung der erkannten Muster mit den Strukturdaten der Website lassen sich typische Nutzerszenarien definieren, die später weiter verwendet werden können. Mittels dieser aggregierten Benutzermodelle lassen sich auch für neue Benutzer Adaptionseffekte erzielen, indem nach kurzer Analyse der Benutzung der jeweilige Benutzer in eines der vordefinierten Stereotypenmodelle eingeordnet wird.

### 4.2.2 Web-2.0-Ansätze zur Benutzermodellierung

Carmagnola et al. (2007b) schlagen zur Erweiterung der Benutzermodellierung im Web-2.0-Umfeld vor, die von einem Benutzer auf verschiedenen interaktiven Websites<sup>4</sup> verwendeten Tags in einer zen-

---

<sup>4</sup>Wie zum Beispiel del.icio.us, flickr.com, youtube.com.

tralen Ontologie zu aggregieren. Aus den Tags, die ein Benutzer zur Beschreibung eines Objektes vergeben hat, lassen sich primär seine Interessen bezüglich mehrerer Bereiche erkennen, darüber hinaus jedoch auch der Interaktionsgrad sowie die Einstellung des Benutzers zur Organisation von Inhalten ableiten. Der Interaktionsgrad bestimmt sich einerseits aus der Bereitschaft des Benutzers, Zeit in das Tagging von Inhalten zu investieren, als auch aus den Möglichkeiten, die ein System dem Benutzer hierfür bietet (Carmagnola et al., 2007a). Um die Tags unterschiedlicher Systeme in einem zentralen Benutzermodell verwenden zu können, werden die Tags aus den individuellen Datenstrukturen der Websites (meist XML-Strukturen) extrahiert, analysiert, kategorisiert und auf das Benutzermodell übertragen. Um die weitere Verwendung der Tags zu ermöglichen, ist es erstrebenswert, die kumulierten Tags anderen Systemen über eine Schnittstelle zur Verfügung zu stellen. Firan et al. (2007) stellen dar, dass Tag-basierte Empfehlungsdienste eine deutlich höhere Qualität als klassische Dienste, die auf kollaborativem Filtern basieren, aufweisen, deutlich schneller arbeiten und keine initialen Kaltstartprobleme besitzen.

Es besteht ein Trend in Richtung zentraler, portabler Benutzerprofile, was auch die Bildung der Arbeitsgruppe zur Schaffung der *Attention Profile Markup Language* (APML)<sup>5</sup> deutlich macht. Mit der APML soll ein Benutzer seine Daten bezüglich seiner Interessen in einem standardisierten Format sammeln und in verschiedenen Diensten weiterverwenden können. Dieser Ansatz verschafft dem Benutzer die Verfügungsgewalt über sein persönliches Interessensprofil, somit kann er selbst entscheiden, welche Daten er weitergeben möchte, und kann sein persönliches Profil jederzeit einsehen und bearbeiten.

---

<sup>5</sup><http://www.apml.org>, zugegriffen am 20.01.2010

### 4.2.3 Datenschutz

Insbesondere im Internet gewinnt das Thema Datenschutz für Benutzer immer mehr an Bedeutung. Kobsa (2007b) bietet eine breite Übersicht über die für personalisierte Systeme relevanten Themen, welche an dieser Stelle überblicksartig dargestellt werden. Während in früheren adaptiven Systemen lediglich eine beschränkte Menge persönlicher Daten gesammelt und meist nur lokal verarbeitet wurde, fließen in adaptiven Web-Anwendungen die Benutzerdaten an einen Web-Server und entziehen sich somit der Kontrolle des Benutzers. Durch häufige bekannt gewordene Sicherheitsbrüche, Hackerangriffe und vermehrte Sammlung persönlicher Daten wächst seitens der Benutzer eine immer stärkere Abneigung gegenüber der Bekanntgabe persönlicher Daten. Technische Fortschritte erlauben es mittlerweile, nicht nur textuelle Informationen über Benutzer zu sammeln, sondern auch Daten wie den Aufenthaltsort, die Mausbewegungen und den Browse-Verlauf durch das Internet in individuelle Benutzermodelle aufzunehmen. Die Aggregation von Daten aus unterschiedlichen Quellen und Analysen über große Datenbestände werden durch stärkere Rechenleistung und größere Bandbreiten erst in jüngster Zeit ermöglicht. Aufgrund dieser Fortschritte werden in vielen Ländern Gesetze zum Datenschutz erlassen, die von Herstellern und Betreibern adaptiver Systeme berücksichtigt werden müssen. Diese Entwicklungen führen zu einer immer stärkeren Beachtung des Datenschutzes im Bereich adaptiver Systeme.

In einer Vielzahl von Studien (Kobsa, 2007b, S. 632) wurde festgestellt, dass zwischen 70% und knapp 90% aller Internetnutzer über ihre Privatsphäre und die Sicherheit ihrer persönlichen Daten im Netz besorgt sind. Zwischen 82% und 95% aller Benutzer vermieden es daher mindestens ein mal, persönliche Daten auf einer Website zu hinterlassen. Knapp ein Viertel der Benutzer gab an, grundsätzlich im Internet keine persönlichen Angaben zu machen. Einige Benutzer neigen dazu, bei der Aufforderung, Daten einzugeben, fiktive oder falsche Informationen zu übermitteln. Fast alle Benutzer haben Bedenken, wenn ein Unternehmen gesammelte per-

sonenbezogene Daten an Drittfirmen übermittelt, die diese zu anderen Zwecken weiter verwenden. Diese Einstellung seitens der Benutzer führt zu deutlichen Problemen beim Aufbau vollständiger und valider Benutzermodelle und verhindert den Aufbau zentraler Benutzermodelle, die in verschiedenen adaptiven Systemen verwendet werden können. Es treffen jedoch nicht nur explizite, sondern auch implizite Methoden der Datensammlung auf Widerstand. Bei der Verfolgung des Browser-Verlaufs mittels Cookies haben mehr als die Hälfte der untersuchten Benutzer Bedenken, insbesondere wenn der Navigationsverlauf über mehrere Websites hinweg nachverfolgt werden kann. Dennoch akzeptieren die meisten Benutzer (62%) generell Cookies, jedoch leeren über die Hälfte der Benutzer gelegentlich den Cookie-Cache ihres Browsers. Dieses Verhalten beeinträchtigt die oben beschriebenen Techniken des Web-Usage-Mining und generell alle adaptiven Web-Systeme.

Die Bereitschaft der Benutzer, persönliche Daten preiszugeben hängt stark von der Art der Informationen ab. Generell ist erkennbar, dass Benutzer bei einfachen demographischen Daten, Angaben ihrer Hobbies oder ihrem Medienkonsumverhalten weniger Bedenken äußern als bei der Angabe von Kreditkarteninformationen, Kontaktinformationen oder finanziellem Status. In der Praxis konnte allerdings eine Diskrepanz zwischen geäußerten Einstellungen bezüglich Datenschutz und tatsächlicher Preisgabe persönlicher Informationen im Internet beobachtet werden. Insbesondere ist die Bereitschaft abhängig vom aktuellen Ziel des Benutzers und den Vorteilen, die sich aus den Angaben für den Benutzer ergeben. Insbesondere im Hinblick auf die sich aus der Personalisierung einer Website ergebenden Vorteile sind Benutzer dazu bereit, persönliche Informationen preiszugeben und erachten es als hilfreich, wenn gewisse Daten gespeichert werden. Positiven Einfluss auf die Einstellung der Benutzer haben insbesondere Informationen über die gesammelten Daten und deren Verwendung, sowie die Möglichkeit der informationellen Selbstbestimmung. Von Bedeutung sind auch das Vertrauen in eine Website, die sich aus positiven Erfahrungen in der Vergangenheit, dem Ruf des Websitebetreibers, der gestalterischen und funktiona-

len Qualität der Website oder Hinweisen auf den Umgang mit dem Thema Datenschutz ergeben.

Kobsa (2007b) beschreibt eine Vielzahl technischer Möglichkeiten, die die Datenschutzprobleme von Benutzermodellen mindern können. Dazu zählt die Möglichkeit, Benutzermodelle auf Pseudonymen basieren zu lassen, die zwar von adaptiven Systemen einem individuellen Benutzer zugeordnet werden können, Drittpersonen allerdings keine Rückschlüsse auf die wahre Identität des Benutzers ermöglichen. Die Speicherung von Benutzermodellen auf der Client-Seite löst Probleme hinsichtlich der Gesetzgebung als auch des Vertrauens des Benutzers. Da keine persönlichen Daten ins Web übertragen werden, sondern auch die Personalisierungskomponente lokal arbeitet, werden Benutzer unter Umständen mehr Daten zur Verfügung stellen. Allerdings schränken client-seitige Techniken die Verwendungsmöglichkeiten, zum Beispiel hinsichtlich kollaborativen Filterns, ein. In der Forschung werden weitere Möglichkeiten untersucht, etwa die verteilte Speicherung von Benutzermodellen, die Arbeit mit verschlüsselten Daten, *Perturbation* (es werden zufällige Werte in das Benutzermodell eingestreut) oder *Obfuscation* (Modifikation gewisser Werte). Anbieter adaptiver Systeme werden also in Zukunft dem Thema Datenschutz immer mehr Bedeutung beimessen müssen.

## 4.3 Inhalte von Benutzermodellen

In diesem Abschnitt werden die möglichen Inhalte von Benutzermodellen dargestellt. In einem konkreten Benutzermodell muss allerdings nicht zwangsläufig jeder Aspekt eines Benutzers hinterlegt werden, vielmehr folgt der Inhalt des Modells letztlich dem gewünschten Einsatzzweck.

### 4.3.1 Wissen und Interessen

Die Modellierung des Wissens eines Benutzers stellt für die meisten Anwendungsbereiche die wichtigste Funktion dar. Besonders in ad-

aptiven Lernumgebungen stellt das Wissen des Benutzers das Kernelement dar. Eine Möglichkeit, das gesamte Wissen eines Benutzers in einem spezifischen Bereich darzustellen, ist die Bewertung auf einer linearen Skala – quantitativ oder qualitativ. Hierbei wird jedoch nicht auf feine Wissensunterschiede in kleineren Teilbereichen eingegangen, sodass nur ein grobgranularer Adaptionseffekt erzielbar ist. Eine Weiterentwicklung stellen strukturelle Modelle dar, die es ermöglichen, das Fachwissen des Benutzers in einer bestimmten Domäne anhand definierter Fragmente darzustellen. Beim sogenannten *Overlay Model* wird das individuelle Wissen des Benutzers in Relation zum Expertenwissen der Domäne bewertet, entweder qualitativ (gut, durchschnittlich, schwach) oder quantitativ (auf einer Skala von 1 bis 5). Da dieses Modellierungsverfahren allerdings keine Informationen über fehlerhaftes Wissen des Benutzers enthält, wurden sogenannte *Bug Models* entwickelt, die hochqualitative Adaption unterstützen. Aufgrund ihrer Komplexität werden *Bug Models* im Gegensatz zu einfacheren *Overlay Models* in der Praxis relativ selten eingesetzt (Brusilovsky und Millán, 2007).

Die Modellierung der Interessen eines Benutzers hat in den letzten Jahren zunehmend an Bedeutung gewonnen. Insbesondere im Bereich von adaptiven Information-Retrieval-Systemen und Filtersystemen, die große Mengen an Informationen verarbeiten müssen, sind Modelle der Interessen von Benutzern von großer Bedeutung. Mit der Verbreitung von adaptiven Informationssystemen wie Enzyklopädien, News-Portalen oder auch E-Commerce-Shops, bei denen die Informationsbeschaffung hauptsächlich auf Interessen des Benutzers basiert, breitet sich die Modellierung der Interessen auf immer größere Bereiche aus und tangiert immer mehr Menschen. Auch bei adaptiven Lernumgebungen wird, forciert durch konstruktivistische Ansätze in der Didaktik, eine interessenbasierte Vermittlung von Lerninhalten verfolgt und damit die Modellierung von Interessen zunehmend wichtiger. Die Abbildung der Interessen in Modellen wurde historisch mittels gewichteter Vektoren von Schlüsselbegriffen durchgeführt, diese Herangehensweise ist heute noch weit verbreitet. Eine detailliertere Abbildung erlauben *Overlay Models*, die ähnlich

aufgebaut sind wie diejenigen zur Repräsentation von Wissen. Hierbei werden spezifische Interessen des Benutzers über Begriffsmodelle bestimmter Domänen gelegt, sodass unterschiedliche Interessensgebiete aus dem Gesamtmodell abgebildet werden können. Durch semantische Verbindungen zwischen einzelnen Begriffen eines Modells können Beziehungen zwischen unterschiedlichen Bereichen eines Interessensgebiets abgebildet werden und somit die Erschließung neuer Gebiete gefördert werden (Brusilovsky und Millán, 2007) (siehe auch Abschnitt 4.4).

### **4.3.2 Aktuelle Ziele und Aufgaben**

Das aktuelle Ziel eines Benutzers herauszufinden, stellt einerseits eine große Herausforderung dar, ist andererseits für die Adaption der Software an die Bedürfnisse des Benutzers von Bedeutung. Dabei kann das Ziel sein, sich Informationen zu beschaffen oder eine Aufgabe zu erledigen, die zur Erfüllung der Arbeit notwendig ist. Das Ziel kann sich sehr schnell verändern, bei jeder erneuten Benutzung einer Software anders sein oder sich sogar während der Benutzung verändern. Abgebildet werden die Ziele heute meist über einen Katalog von möglichen Zielen, die eine Software erkennen und unterstützen kann. Durch einen hierarchischen Aufbau und Verknüpfungen können Abhängigkeiten definiert werden, die bei der Adaption beachtet werden müssen. Ein Modellierungssystem muss also das aktuelle Ziel des Benutzers erkennen können und dieses in die bestehende Zielhierarchie einordnen. Basierend auf dieser Einordnung werden dann Adaptionsmechanismen gestartet, beispielsweise die Navigation angepasst oder passende Inhalte präsentiert, die bei der Erfüllung des Ziels hilfreich sind.

Um die Ziele eines Benutzers definieren zu können, bestehen unterschiedliche Herangehensweisen. Die einfachste Variante ist es, den Benutzer selbst aktiv aus einem Katalog von Zielen auswählen zu lassen. Die Berechnung der wahrscheinlichsten Ziele im Verlauf der Benutzung stellt eine andere Möglichkeit dar. Hierbei kann allerdings nur mit Zeitverzögerung ein Adaptionseffekt erzielt werden.

Die aktuellste Möglichkeit stellt die Verwendung von Data-Mining-Techniken dar, um aus dem Nutzungsverlauf die wahrscheinlichsten nächsten Ziele ableiten zu können. So lassen sich bereits ab Beginn der Benutzung im aktuellen Nutzungskontext Adaptionen des Systems durchführen. In der Regel erlauben automatisierte Verfahren auch den manuellen Eingriff des Benutzers, um sein aktuelles Ziel präzise zu definieren (Brusilovsky und Millán, 2007).

### 4.3.3 Hintergrund, Charakterzüge und Arbeitskontext

Der persönliche Hintergrund des Benutzers umfasst Eigenschaften wie Erfahrung außerhalb des Fachgebiets des Systems, den Beruf, Erfahrungen in verwandten Bereichen, demographische Daten oder auch die Meinung über gewisse Sachverhalte. Aus diesen Angaben lassen sich Sachverhalte ableiten, die für die Adaption notwendig sind. So lässt sich beispielsweise anhand des Berufes des Benutzers dessen Kenntnisstand von Fachbegriffen ableiten und die Sprache des Inhaltes daran anpassen. Diese Hintergrundinformationen verändern sich während der Benutzung kaum oder überhaupt nicht, daher ist meist keine individuelle Modellierung notwendig, sondern es kann auf Stereotypenmodelle zurückgegriffen werden. Im Gegensatz zu Wissen können Hintergrundinformationen zudem kaum aus der Benutzungsweise eines Systems abgeleitet werden, sodass diese Angaben explizit vom Benutzer eingegeben werden müssen.

Charaktereigenschaften können nicht aus der Benutzung abgeleitet werden, sondern müssen über psychologische Tests festgestellt werden. Dazu zählen beispielsweise Temperament, Denkstil und andere kognitive Faktoren wie der persönliche Lernstil. Insbesondere an den kognitiven Stil des individuellen Benutzers lassen sich Anpassungen durchführen. Beispielsweise können die Navigation durch das System, Anpassungsmöglichkeiten des Benutzers oder Hilfestellungen personalisiert werden. Der Nutzen einer derartigen Anpassung ist evident, allerdings sind derzeit noch keine verlässlichen Umsetzungen adaptiver Systeme, die sich an kognitive Faktoren anpassen, verfügbar.

Die Adaption an den Arbeitskontext ist insbesondere mit dem Aufkommen von Web-basierten Systemen bedeutend geworden. Auch zunehmend mobile Zugriffe auf Web-basierte Systeme erfordern eine Anpassung an spezifische Anforderungen der Benutzer dieser Geräte. So muss sich eine Anwendung an die Plattform des Benutzers anpassen können, da diese unterschiedliche technische Voraussetzungen bietet, wie beispielsweise Bildschirmauflösung, Bandbreite oder Eingabemöglichkeiten. Insbesondere im Mobile-Bereich wird auch die Erfassung des Standortes des Benutzers zur Anpassung immer bedeutender, da hier technische Lösungen die Darstellung standortbezogener Informationen ermöglichen (Brusilovsky und Millán, 2007).

## 4.4 Formale Darstellung von Benutzermodellen

Die Darstellung von Benutzermodellen wird in diesem Abschnitt beschrieben. Dabei wird zunächst auf stereotypisierte Modelle, im Anschluss auf individuelle Modelle mit ihren unterschiedlichen Darstellungsformen eingegangen. Anhand einer beispielhaften Implementierung wird das Konzept von Ontologien zur Benutzermodellierung vorgestellt. Abschließend wird auf die Speicherung von Benutzermodellen eingegangen.

### 4.4.1 Stereotype und individuelle Modelle

Die Benutzung von Stereotypen zur Benutzermodellierung geht auf Rich (1979, 1983) zurück. In dieser Zeit wurden Benutzermodelle primär in Dialogsystemen eingesetzt, die aus expliziten Angaben des Benutzers gebildet wurden. Um die Zahl der notwendigen Angaben des Benutzers möglichst zu minimieren, soll hier der Benutzer nach kurzer Zeit in eine Benutzergruppe eingeordnet werden, die eine Vielzahl der möglichen Attribute des Benutzers abbildet. Stereotypisierte Modelle müssen also in einem Modellierungssystem verfügbar sein, es muss eine Möglichkeit gegeben sein, den Benutzer zuzuordnen und im Lauf der Benutzung in immer feinere Stereotypen einzuordnen. Jeder Stereotyp besitzt eine Menge von Attributen, die

einzelnen gewichtet sind und deren Übereinstimmungsgrad mit den Informationen über den Benutzer eine Wahrscheinlichkeit beschreibt, mit der der Nutzer diesem Typus zugeordnet werden kann. Zudem ist in dem Stereotyp ein *Trigger* definiert, der die Aktivierung dieses Stereotypen auslöst. Alle Stereotypen in einem System werden in einer Hierarchie dargestellt, die es ermöglicht, das Modell des Benutzers in eine immer feinere Kategorie einzuordnen.

Stereotypen bilden in der Regel lediglich eine Grundlage zur Erstellung individueller Benutzermodelle, da sie helfen, anfängliche Kaltstartprobleme zu überwinden, die auftreten, wenn keine ausreichende Menge an Informationen über den Nutzer verfügbar ist, um einen präzisen Adaptionseffekt zu erzielen. Für eine präzise Anpassung ist eine individuelle Benutzermodellierung allerdings meist zwingend erforderlich.

#### Design

Interfacedesign	Typografie	Produktdesign	Modedesign	Webdesign	...
0,67	0,90	0,40	0,10	0,30	...

#### Literatur

Dramatik	Fantastik	Horror	Krimi	Lyrik	...
0,95	0,31	0,10	0,88	0,28	...

#### Musik

Alternative	Hip-Hop	Funk	Klassik	Ska	...
0,78	0,74	0,27	0,45	0,10	...

**Abbildung 4.3:** Ein auf Schlagworten basierendes Benutzermodell.

### 4.4.2 Benutzerindividuelle Schlagwort-Profile

Benutzermodelle, die aus einer Menge von Schlagwörtern bestehen, werden als *Keyword Profiles* bezeichnet (Gauch et al., 2007). In Abbildung 4.3 ist ein solches Profil beispielhaft dargestellt. In der Re-

gel wird jedes Schlagwort mit einer Gewichtung versehen, die den Interessens- oder Kenntnisgrad des Nutzers im jeweiligen Bereich repräsentiert. Der Detailgrad der Schlagworte kann oberflächlich oder sehr tiefgehend gehalten werden. Durch Kategorisierung der einzelnen Worte lassen sich Interessensschwerpunkte des Benutzers ausmachen. Die Schlagworte können beispielsweise aus Dokumenten, beispielsweise Webseiten, die ein Benutzer betrachtet und positiv bewertet, mittels Text-Mining-Techniken extrahiert und gewichtet werden.

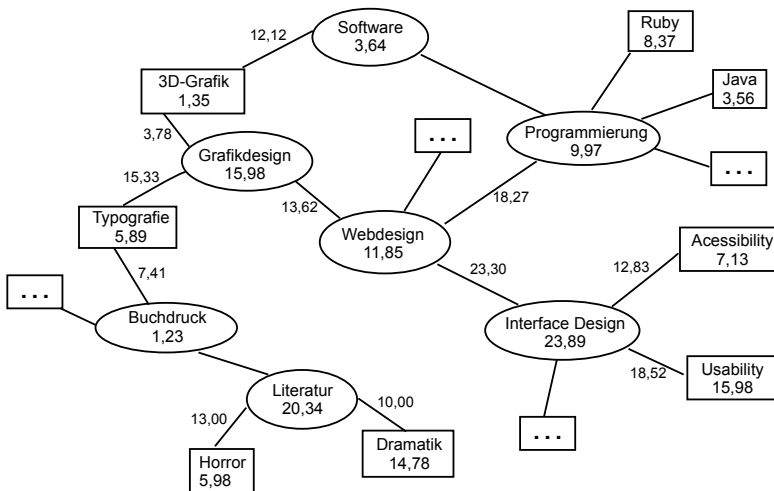


Abbildung 4.4: Benutzermodell auf Basis eines semantischen Netzes

### 4.4.3 Profile basierend auf Semantischen Netzen

Eine erweiterte Form der Schlagwort-Profile stellen Profile basierend auf Semantischen Netzen (*Semantic Network Profiles*) dar (siehe

Abbildung 4.4). Hierbei werden zusätzlich Verbindungen zwischen einzelnen Schlagworten abgebildet und diese Verbindungen ebenfalls mit einer Gewichtung versehen. Die Knoten in diesen Netzen können übergeordnete Begriffe oder Klassen darstellen (Hyperonyme), deren assoziierte Begriffe oder Unterbegriffe (Hyponyme) mittels Kanten verbunden sind. In einfachen Varianten können die Knoten aus konkreten Begriffen ohne Bezug zu ihren jeweiligen Hyperonymen bestehen.

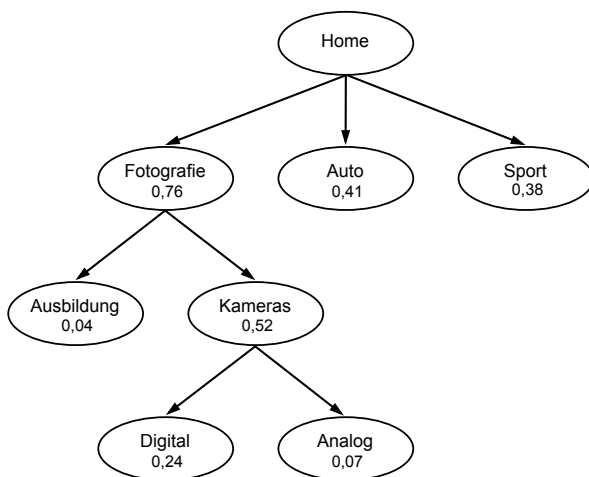
Zur Konstruktion komplexer Semantischer Netze mit Bezug zu Hyperonymen muss eine Möglichkeit geschaffen werden, neue Begriffe in ein bestehendes Netz einzusortieren, hierzu kann auf Worthierarchien wie WordNet<sup>6</sup> zurückgegriffen, eigene Lernmechanismen implementiert oder eine manuelle Sortierung vorgenommen werden. Bei einfachen Varianten werden lediglich Begriffe, die gleichzeitig in Dokumenten vorkommen, miteinander verbunden. Zur Darstellung der Benutzerinteressen wird für jeden Benutzer ein eigenständiges Semantisches Netz aufgebaut. Analog zu Schlagwort-Profilen sind die Knoten je nach Interessengrad des Benutzers gewichtet, darüber hinaus stellen gewichtete ungerichtete Graphen die Stärke der Affinität zwischen den Begriffen dar (Gauch et al., 2007).

#### 4.4.4 Profile basierend auf Begriffshierarchien

Begriffsbasierte Profile (*Concept Profiles*) sind Profilen, die auf Semantischen Netzen basieren, sehr ähnlich, da auch sie aus Knoten bestehen, die über Kanten miteinander verbunden sind. Der Unterschied liegt darin, dass die Knoten aus abstrakten Themen bestehen, nicht aus konkreten Begriffen oder Ausprägungen. Auch wird bei der Bildung des Nutzerprofils immer auf bestehenden Begriffshierarchien aufgebaut und kein individuelles Begriffsnetzwerk aufgebaut. Die einzelnen Knoten werden je nach Wissens- oder Interessensgrad gewichtet, wie es auch bei Schlagwort-Profilen der Fall ist (siehe Abbildung 4.5).

---

<sup>6</sup><http://wordnet.princeton.edu>, zugegriffen am 01.02.2010



**Abbildung 4.5:** Ein auf einer Begriffshierarchie basierendes Benutzermodell.

Zur Bildung des Profils wird häufig auf Verzeichnisse wie das *Open Directory Project*<sup>7</sup> oder die ACM-Klassifikation<sup>8</sup> zurückgegriffen. Mittels impliziter oder expliziter Techniken werden beispielsweise besuchte Webseiten analysiert und in eine Hierarchieebene eingestuft. Durch die Verwendung eines hierarchischen Profils können Generalisierungen durchgeführt werden, welche die Erweiterung des Interessengebietes und die Einbeziehung verwandter Bereiche ermöglichen. In dieser einfachen Form von begriffsbasierten Benutzermodellen besitzen die einzelnen Begriffe maximal *ist-ein* oder *hat-ein* Beziehungen zueinander. Eine weiter entwickelte Form stellen Ontologien dar, in welchen weitere Beziehungstypen abgebildet werden können (Gauch et al., 2007).

---

<sup>7</sup><http://www.dmoz.org>, zugegriffen am 20.01.2010

<sup>8</sup><http://www.acm.org/about/class/ccs98-html>, zugegriffen am 20.01.2010

#### 4.4.5 Datenrepräsentation durch Ontologien

In der Informatik bezeichnet der Begriff *Ontologie* eine formale Beschreibung von Begriffen und deren Beziehungen in einem bestimmten Fachgebiet (Antoniou und van Harmelen, 2008; Sieg et al., 2007). Die wichtigsten Auszeichnungssprachen zur Repräsentation von Ontologien umfassen das *Resource Description Framework*<sup>9</sup> (RDF) als Datenmodell für Objekte und deren Beziehungen bzw. *RDF Schema* (RDFS) zur Beschreibung von Klassen von RDF-Objekten und die *Web Ontology Language*<sup>10</sup> (OWL). Während mit RDF lediglich Super- bzw. Subklassenbeziehungen abgebildet werden können, bietet OWL deutlich umfassendere Möglichkeiten, beispielsweise zur logischen Repräsentation von Beziehungen, Einschränkungen oder Disjunktionen. OWL baut technisch auf RDF und RDFS auf, basiert also ebenfalls auf der *Extensible Markup Language (XML)* (siehe Antoniou und van Harmelen (2008) für weitere Details zur OWL).

Die Verwendung von Ontologien und insbesondere der OWL zur Abbildung von Benutzermodellen bringt einige Vorteile mit sich. Anhand der ontologischen Darstellung des Benutzermodells und einem Domänenmodell lassen sich Inferenzen durchführen, wodurch beispielsweise unbekanntes Wissen über den Benutzer entdeckt werden kann. Die Verwendung einer formalisierten Sprache erleichtert die Weiterverwendung von Benutzermodellen in unterschiedlichen Anwendungen sowie die zentrale Verwaltung und Pflege der Benutzermodelle. Es bestehen bereits Ansätze, um eine allgemeingültige Benutzer-Ontologie zu schaffen, die von vielen Diensten gemeinsam verwendet werden kann (Andrejko et al., 2006).

Heckmann et al. (2005a,b) schlagen eine auf OWL basierende generische Ontologie für Benutzermodelle vor, die *General User Model Ontology* (GUMO), die eine Mehrfachverwendung von Benutzermodellen durch einen einheitlichen Standard begünstigen soll. GUMO deckt einen sehr großen Bereich möglicher Attribute eines Benutzers ab, wie beispielsweise grundlegende persönliche Informationen

---

<sup>9</sup><http://www.w3.org/RDF>, zugegriffen am 20.01.2010

<sup>10</sup><http://www.w3.org/2004/OWL>, zugegriffen am 20.01.2010

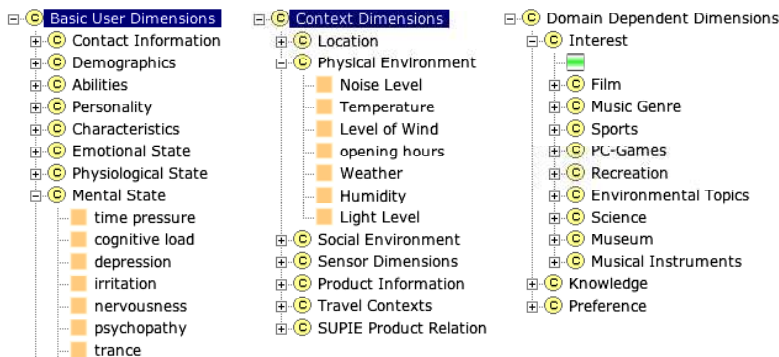


Abbildung 4.6: Ausschnitt der GUMO (Quelle: Ubisworld (2009)).

(z.B. Demographie, Charakter, Fähigkeiten), domänenabhängige Informationen wie verschiedene Interessens- und Wissensgebiete oder auch physikalische Daten aus seiner Umgebung (z.B. Temperatur, Luftfeuchtigkeit oder Wetterbedingungen) (siehe Abbildung 4.6).

Einzelne Attribute werden als OWL-Klassen abgebildet, die je nach Typ unterschiedliche Eigenschaften haben können. Beispielfhaft ist im Folgenden das Attribut der kognitiven Auslastung als XML-Code aufgeführt (Heckmann, 2006).

```
<owl:Class rdf:ID="CognitiveLoad.800102">
  <rdfs:label> cognitive load </rdfs:label>
  <ubis:identifizier> 800102 </ubis:identifizier>
  <ubis:category> 45 </ubis:category>
  <dc:creator rdf:resource="#ubisworld;subject=210003" />
  <ubis:durability> Hour.520060 </ubis:durability>
  <ubis:website rdf:resource="#ubisworld;subject=800102" />
  <rdfs:subClassOf rdf:resource="#MentalState.700017" />
</owl:Class>
```

Da dieses Beispiel aus der Ubis-Ontologie stammt, wird deren Namespace *ubis* verwendet. Erkennbar ist an diesem Beispiel, dass jede Klasse eine eindeutige Identifikationsnummer besitzt, in eine

Kategorie eingeordnet wird und eine Subklassenbeziehung besitzt. Bei dieser Klasse wird außerdem die Lebensdauer des Attributes definiert, da sich die Werte stündlich ändern können. In Heckmann (2005) findet sich eine sehr umfassende Beschreibung der GUMO, welche auch ein repräsentatives Beispiel der Inhalte von Benutzermodellen darstellt. Eine domänenspezifischere Ontologie beschreiben Andrejko et al. (2006), hierbei heben die Autoren auch die Vorteile ontologischer Benutzermodelle hervor. Dolog und Nejdil (2007) beschreiben weitere Einsatzmöglichkeiten von Techniken des *Semantic Web* im Bereich adaptiver Systeme und gehen dabei auch auf die Möglichkeiten des *Reasoning* in Ontologien ein. Auch in verwandten Bereichen wie dem Usage Mining können Techniken des Semantic Web, wie zum Beispiel Ontologien, eingesetzt werden (Stumme et al., 2002).

#### 4.4.6 Speicherung von Benutzermodellen

Die Verwaltung der Benutzermodelle geschieht heute in der Regel über spezielle Benutzermodellierungsserver (*User Modeling Servers*). Die zentrale Speicherung bringt den Vorteil mit sich, dass Benutzermodelle über mehrere Anwendungen hinweg verwendet und auch aus verschiedenen Anwendungen mit neuen Informationen versorgt werden können. Dadurch werden Redundanzen vermieden und die Datenkonsistenz kann einfacher sichergestellt werden. Durch die Aggregation mehrerer Modelle können Stereotypen und Benutzergruppen gebildet und gepflegt werden, wie sie in adaptiven Systemen häufig eingesetzt werden. Zudem müssen personenbezogene Daten mit deutlich größerer Sorgfalt behandelt werden, auch um Datenschutzgesetzen zu genügen. Zu diesem Zwecke kann in zentralen Servern der Zugriff, Sicherheit und Verschlüsselung der Daten verwaltet werden (Kay et al., 2002; Kobsa, 2001, 2007a).

Es existieren vielfältige Anforderungen an generische Benutzermodellierungssysteme, einen Überblick bietet Kobsa (2007a). Zu den für Web-Anwendungen wichtigsten Aspekten zählen:

- Die Unterstützung einer schnellen Adaption, die je nach Verfügbarkeit von Benutzerdaten eine entsprechende Personalisierung bietet.
- Die Erweiterbarkeit um neue Personalisierungsmethoden und Datenquellen und die Einbindung externer Systeme.
- Die Integration externer Datenquellen wie Kundendatenbanken oder Marketing-Daten, da diese beim Start einer personalisierten Anwendung in einem Unternehmen meist schon vorliegen. Diese Daten müssen auch weiterhin häufig parallel gepflegt werden können.
- Die Verwaltung verteilter Informationen wird in Zukunft von immer größerer Bedeutung sein, da Daten aus verschiedenen Quellen gesammelt werden und sich diese immer weiter ausbreiten, wie zum Beispiel auf mobile Endgeräte.
- Die Unterstützung offener Standards sowohl zur Speicherung von Benutzermodellen als auch zum Zugriff auf diese (beispielsweise mittels des *Lightweight Directory Access Protocol* (LDAP)) ist für die Sicherung der zukünftigen Erweiterung und Interoperabilität von Bedeutung.
- Da das Zugriffsvolumen bei Web-Anwendungen stark schwanken kann, müssen Techniken des Load-Balancing und Notlösungen bei möglichen Ausfällen implementiert werden.
- Wenn Benutzermodelle aus unterschiedlichen Quellen gebildet werden, muss eine Transaktionssicherheit gegeben sein, die die Datenkonsistenz bei gleichzeitigen Schreib- oder Lesevorgängen erhält.
- Aspekte des Datenschutzes werden immer wichtiger, daher muss das Benutzermodellierungssystem diesbezüglich vielfältige Unterstützung bieten.

Des Weiteren sind Benutzermodellierungsserver in der Lage, Annahmen über noch unbekannte Eigenschaften des Benutzers zu berechnen. Sie sammeln Informationen über die Benutzung des Systems und leiten daraus Annahmen ab oder ermöglichen die Definition von Regeln, um neue Benutzer bestimmten Gruppen zuzuweisen (Kobsa, 2007a).

## 4.5 Zusammenfassung und Ausblick

In diesem Kapitel wurde ein Überblick über das Thema Benutzermodellierung in adaptiven Web-Anwendungen gegeben. In einer Einführung wurden die wichtigsten Begriffe erläutert. Es wurde auf verschiedene Techniken zur Gewinnung von Benutzer- und Nutzungsdaten eingegangen und mit dem Web-Usage-Mining ein aktueller Ansatz der Datengewinnung im WWW vorgestellt. Das Thema Datenschutz, welches bei der Verarbeitung personenbezogener Daten von Bedeutung ist, wurde im Zusammenhang mit der Benutzermodellierung besprochen. Als mögliche Inhalte von Benutzermodellen wurden unter anderem das Wissen und die Interessen, die Ziele und der Arbeitskontext des Benutzers identifiziert. Unterschiedliche formale Darstellungsmöglichkeiten von Benutzermodellen wurden anschließend besprochen und die GUMO als mögliche Grundlage Ontologie-basierter Benutzermodelle exemplarisch vorgestellt. Die wichtigsten Aspekte von Serversystemen zur Verwaltung von Benutzermodellen wurden ebenfalls in diesem Abschnitt dargelegt.

In Zukunft wird das Thema Benutzermodellierung in weitere Bereiche außerhalb klassischer Computersysteme vordringen. Durch Fortschritte der Mobiltechnik werden Einsatzszenarien, die eine ubiquitäre Benutzermodellierung (Heckmann, 2005) mittels Mobiltelefonen oder ähnlichen Geräten beinhalten, immer realistischer. Gleichzeitig weiten sich die Einsatzbereiche von Benutzermodellen zur Personalisierung auf Geräte der realen Welt aus. So existieren beispielsweise elektronische Autoschlüssel, die Sitzposition, Spiegelstellung und Navigationssystem auf den jeweiligen Fahrer anpassen. Je mehr

Geräte kleine Benutzermodelle besitzen, desto stärker wird die Standardisierung und Zentralisierung von Benutzermodellen gefördert (Kobsa, 2007a). Mit wachsenden Benutzermodellen, vermehrter Weiterverwendung und Integration von Daten aus unterschiedlichen Quellen gewinnt der Datenschutz besonders für die Benutzer immer mehr an Bedeutung. Benutzer und Legislatur werden diesbezüglich auch in Zukunft wachsende Anforderungen an die Hersteller und Anbieter adaptiver Web-Anwendungen stellen.

## Literatur

- Andrejko, A., Barla, M., und Bielikov, M. (2006). Ontology-based user modeling for web-based information systems. In: Maygar, G., Knapp, G., Wojtkowski, W., Wojtkowski, G., und Zupancic, J. (Hrsg.), *Advances in Information Systems Development: New Methods and Practice for the Networked Society*, Band 2, Berlin. Springer.
- Antoniou, G. und van Harmelen, F. (2008). *A Semantic Web Primer, second edition*. The MIT Press.
- Baumeister, J. und Atzmüller, M. (Hrsg.) (2008). *LWA 2008 - Workshop-Woche: Lernen, Wissen & Adaptivität, Würzburg, 6.-8. Oktober 2008, Proceedings*, Band 448 in *Technical Report*. Department of Computer Science, University of Würzburg, Germany.
- Berendt, B., Hotho, A., und Stumme, G. (2002). Towards Semantic Web Mining. In: Horrocks, I. und Hendler, J. A. (Hrsg.), *ISWC '02: Proceedings of the First International Semantic Web Conference on The Semantic Web*, S. 264–278, London, UK. Springer-Verlag.
- Berkovsky, S., Kuflik, T., und Ricci, F. (2009). Cross-representation mediation of user models. *User Modeling and User-Adapted Interaction*, 19(1 - 2):35–63.
- Brusilovsky, P. (1996). Methods and techniques of adaptive hypermedia. *User Modeling and User-Adapted Interaction*, 6(2):87–129.
- Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, 11(1 - 2):87–110.
- Brusilovsky, P., Kobsa, A., und Nejdl, W. (Hrsg.) (2007). *The Adaptive Web: Methods and Strategies of Web Personalization*, Band 4321 in *Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg.

- Brusilovsky, P. und Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. In Brusilovsky et al. (2007), Kapitel 1, S. 3–53.
- Carmagnola, F., Cena, F., Console, L., Cortassa, O., Gena, C., Goy, A., Torre, I., Toso, A., und Venero, F. (2008). Tag-based user modeling for social multi-device adaptive guides. *User Modeling and User-Adapted Interaction*, 18(5):497–538.
- Carmagnola, F., Cena, F., Cortassa, O., Gena, C., und Torre, I. (2007a). Towards a Tag-Based User Model: How Can User Model Benefit from Tags? In: Conati, C., McCoy, K., und Paliouras, G. (Hrsg.), *User Modeling 2007 - 11th International Conference, UM 2007, Corfu, Greece, July 25-29, 2007, Proceedings*, Band 4511 in *Lecture Notes in Computer Science*, S. 445–449, Berlin, Heidelberg. Springer.
- Carmagnola, F., Cena, F., und Gena, C. (2007b). User Modeling in the Social Web. In: Apolloni, B., Howlett, R. J., und Jain, L. C. (Hrsg.), *Knowledge-Based Intelligent Information and Engineering Systems*, Band 4694 in *Lecture Notes in Computer Science*, S. 745–752, Berlin, Heidelberg. Springer.
- Cooley, R., Mobasher, B., und Srivastava, J. (1997). Web mining: information and pattern discovery on the World Wide Web. In: *Ninth IEEE International Conference on Tools with Artificial Intelligence, 1997. Proceedings.*, S. 558–567.
- De Bra, P., Aerts, A., Berden, B., de Lange, B., Rousseau, B., Santic, T., Smits, D., und Stash, N. (1998). AHA! The Adaptive Hypermedia Architecture. *The New Review of Hypermedia and Multimedia*, 4:115–139.
- Dolog, P. und Nejdl, W. (2007). Semantic Web Technologies for the Adaptive Web. In Brusilovsky et al. (2007), Kapitel 23, S. 697–719.
- Fink, J. und Kobsa, A. (2000). A Review and Analysis of Commercial User Modeling Servers for Personalization on the World Wide Web. *User Modeling and User-Adapted Interaction*, 10(2-3):209–249.
- Firan, C. S., Nejdl, W., und Paiu, R. (2007). The Benefit of Using Tag-Based Profiles. In: Almeida, V. A. F. und Baeza-Yates, R. A. (Hrsg.), *Fifth Latin American Web Congress (LA-Web 2007), 31 October - 2 November 2007, Santiago de Chile*, S. 32–41. IEEE Computer Society.
- Fischer, G. (2001). User Modeling in Human-Computer Interaction. *User Modeling and User-Adapted Interaction*, 11(1):65–86.

- Gauch, S., Speretta, M., Chandramouli, A., und Micarelli, A. (2007). User Profiles for Personalized Information Access. In Brusilovsky et al. (2007), Kapitel 2, S. 54–89.
- Heckmann, D. (2003). Introducing Situational Statements as an integrating Data Structure for User Modeling, Context-Awareness and Resource-Adaptive Computing. In: Hoto, A. und Stumme, G. (Hrsg.), *LLWA Lehren - Lernen - Wissen - Adaptivität (ABIS2003)*, Karlsruhe, Germany.
- Heckmann, D. (2005). *Ubiquitous User Modeling*. PhD thesis, Universität des Saarlandes, Saarbrücken.
- Heckmann, D. (2006). GUMO. <http://www.gumo.org/ubisworld/documents/-gumo/2.0/gumo.owl>. Letzter Zugriff am 06.08.2009.
- Heckmann, D., Schwartz, T., Brandherm, B., und Kröner, A. (2005a). Decentralized User Modeling with UserML and GUMO. In: Dolog, P. und Vassileva, J. (Hrsg.), *Decentralized, Agent Based and Social Approaches to User Modeling, Workshop DASUM-05 at 9th International Conference on User Modelling, UM2005*, S. 61–66, Edinburgh, Scotland.
- Heckmann, D., Schwartz, T., Brandherm, B., Schmitz, M., und von Wilamowitz-Moellendorf, M. (2005b). GUMO The General User Model Ontology. In: Ardissono, L., Brna, P., und Mitrovic, A. (Hrsg.), *User Modeling 2005: 10th International Conference, UM 2005, Edinburgh, Scotland, UK, July 24-29, 2005, Proceedings*, Band 3538 in *Lecture Notes in Computer Science*, S. 428–432. Springer, Berlin, Heidelberg.
- Henze, N. (2000). *Adaptive Hyperbooks: Adaptation for Project-Based Learning Resources*. PhD thesis, Universität Hannover.
- Herder, E. (2006). *Forward, back and home again : analyzing user behavior on the web*. PhD thesis, University of Twente.
- Hinneburg, A. (Hrsg.) (2007). *LWA 2007: Lernen - Wissen - Adaption, Halle, September 2007, Workshop Proceedings*. Martin-Luther-University Halle-Wittenberg.
- Jameson, A. (2003). Adaptive interfaces and agents. In: Jacko, J. A. und Sears, A. (Hrsg.), *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*, Kapitel 15, S. 305–330. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.
- Johnson, A. und Taatgen, N. (2005). User Modeling. In: Proctor, R. W. und Vu, K.-P. L. (Hrsg.), *The Handbook of Human Factors in Web Design*, Kapitel 25, S. 424–438. Erlbaum.

- Kay, J., Kummerfeld, B., und Lauder, P. (2002). Personis: A Server for User Models. In: De Bra, P., Brusilovsky, P., und Conejo, R. (Hrsg.), *Adaptive Hypermedia and Adaptive Web-Based Systems: Second International Conference, AH 2002 Málaga, Spain, May 29-31, 2002 Proceedings*, Band 2347 in *Lecture Notes in Computer Science*, S. 203–212. Springer, Berlin, Heidelberg.
- Kobsa, A. (1985). *Benutzermodellierung in Dialogsystemen*, Band 115 in *Informatik-Fachberichte*. Springer, Berlin, Heidelberg.
- Kobsa, A. (1993). Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen. In: O. Herzog, T. C. u. D. S. (Hrsg.), *17. Fachtagung KI*, Berlin, Heidelberg. Springer.
- Kobsa, A. (2001). Generic User Modeling Systems. *User Modeling and User-Adapted Interaction*, 11(1):49–63.
- Kobsa, A. (2004). Adaptive Verfahren Benutzermodellierung. In: R. Kuhlen, T. S. und Strauch, D. (Hrsg.), *Grundlagen der Information und Dokumentation*, S. 299–302. K. G. Saur, Munich, 5. Auflage.
- Kobsa, A. (2007a). Generic User Modeling Systems. In Brusilovsky et al. (2007), Kapitel 4, S. 136–154.
- Kobsa, A. (2007b). Privacy-Enhanced Web Personalization. In Brusilovsky et al. (2007), Kapitel 21, S. 628–670.
- Langley, P. (1999). User modeling in adaptive interfaces. In: Kay, J. (Hrsg.), *UM '99: Proceedings of the seventh international conference on User modeling*, S. 357–370, Secaucus, NJ, USA. Springer-Verlag New York, Inc.
- Mobasher, B. (2004). Web Usage Mining and Personalization. In: Singh, M. P. (Hrsg.), *Practical Handbook of Internet Computing*. Chapman Hall & CRC Press, Baton Rouge.
- Mobasher, B. (2006). Web Usage Mining. In: Liu, B. (Hrsg.), *Web Data Mining: Exploring Hyperlinks, Contents and Usage Data*, Data-Centric Systems and Applications, S. 449–483. Springer, Berlin, Heidelberg.
- Mobasher, B. (2007). Data Mining for Web Personalization. In Brusilovsky et al. (2007), S. 90–135.
- Mobasher, B., Cooley, R., und Srivastava, J. (2000). Automatic personalization based on Web usage mining. *Communications of the ACM*, 43(8):142–151.

- Omari, A. und Conrad, S. (2006). Web Usage Mining for Adaptive and Personalized Websites. In: Althoff, K.-D. und Schaaf, M. (Hrsg.), *LWA 2006: Lernen - Wissensentdeckung - Adaptivität, Hildesheim, October 9th-11th 2006, joint workshop event of several interest groups of the German Society for Informatics (GI) - 14th Workshop on Adaptivity and User Modeling in Interactive Systems (ABIS 2006) - Workshop Information Retrieval 2006 of the Special Interest Group Information Retrieval (FGIR 2006) - Workshop on Knowledge and Experience Management (FGWM 2006) - 12th Workshop on Knowledge Discovery, Data Mining, and Machine Learning (KDML 2006)*, Hildesheimer Informatik-Berichte. University of Hildesheim, Institute of Computer Science.
- Pazzani, M. und Billsus, D. (2007). Content-Based Recommendation Systems. In Brusilovsky et al. (2007), Kapitel 10, S. 325–341.
- Pierrakos, D., Paliouras, G., Papatheodorou, C., und Spyropoulos, C. D. (2003). Web Usage Mining as a Tool for Personalization: A Survey. *User Modeling and User-Adapted Interaction*, 13(4):311–372.
- Rich, E. (1979). User modeling via stereotypes. *Cognitive Science*, 3(4):329–354.
- Rich, E. (1983). Users are individuals: individualizing user models. *International Journal of Man-Machine Studies*, 18:199–214.
- Sarwar, B., Karypis, G., Konstan, J., und Reidl, J. (2001). Item-based collaborative filtering recommendation algorithms. In: *WWW '01: Proceedings of the 10th international conference on World Wide Web*, S. 285–295, New York, NY, USA. ACM.
- Schafer, J., Frankowski, D., Herlocker, J., und Sen, S. (2007). Collaborative Filtering Recommender Systems. In Brusilovsky et al. (2007), Kapitel 9, S. 291–324.
- Sharma, A. (2001). A Generic Architecture for User Modeling Systems and Adaptive Web Services. In: Nebel, B. (Hrsg.), *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence, IJCAI 2001, Workshop on E-Business & the Intelligent Web*, Seattle, USA. Morgan Kaufmann.
- Shen, X., Tan, B., und Zhai, C. (2005). Implicit user modeling for personalized search. In: Herzog, O., Schek, H.-J., Fuhr, N., Chowdhury, A., und Teiken, W. (Hrsg.), *Proceedings of the 2005 ACM CIKM International Conference on Information and Knowledge Management, Bremen, Germany, October 31 - November 5, 2005*, S. 824–831, New York, NY, USA. ACM Press.

- Sieg, A., Mobasher, B., und Burke, R. (2007). Web search personalization with ontological user profiles. In: Silva, M. J., Laender, A. H. F., Baeza-Yates, R. A., McGuinness, D. L., Olstad, B., Olsen, Ø. H., und Falcão, A. O. (Hrsg.), *Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management, CIKM 2007, Lisbon, Portugal, November 6-10, 2007*, S. 525–534, New York, NY, USA. ACM.
- Smyth, B. und Cotter, P. (2002). Personalized Adaptive Navigation for Mobile Portals. In: van Harmelen, F. (Hrsg.), *Proceedings of the 15th European Conference on Artificial Intelligence, ECAI'2002, Lyon, France, July 2002*, S. 608–612. IOS Press.
- Stumme, G., Hotho, A., und Berendt, B. (2002). Usage Mining for and on the Semantic Web. In: *National Science Foundation Workshop on Next Generation Data Mining*, Baltimore, USA.
- Sugiyama, K., Hatano, K., und Yoshikawa, M. (2004). Adaptive web search based on user profile constructed without any effort from users. In: Feldman, S. I., Uretsky, M., Najork, M., und Wills, C. E. (Hrsg.), *Proceedings of the 13th international conference on World Wide Web, WWW 2004, New York, NY, USA, May 17-20, 2004*, S. 675–684. ACM Press.
- Ubisworld (2009). <http://ontology.ubisworld.org>. Letzter Zugriff am 06.08.2009.
- Wahlster, W. und Kobsa, A. (1989). User Models in Dialog Systems. In: Kobsa, A. und Wahlster, W. (Hrsg.), *User Models in Dialog Systems*, S. 4–34, Berlin. Springer.
- Zhu, T. und Greiner, R. (2003). Learning a model of a web users interests. In: Brusilovsky, P., Corbett, A. T., und de Rosis, F. (Hrsg.), *User Modeling 2003, 9th International Conference, UM 2003, Johnstown, PA, USA, June 22-26, 2003, Proceedings*, Band 2702 in *Lecture Notes in Computer Science*, S. 65–75. Springer.