



Hochschule für Angewandte Wissenschaften Hamburg  
*Hamburg University of Applied Sciences*

## **Ausarbeitung Seminar (AI)**

Thomas Schmidt

Automatische Erstellung eines Benutzerprofils für Social  
Navigation

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Grundlagen</b>	<b>2</b>
2.1	Social Navigation . . . . .	2
2.2	Benutzerprofil . . . . .	2
<b>3</b>	<b>Zwischenmenschliche Strukturen und Personen</b>	<b>3</b>
3.1	Reale Welt . . . . .	3
3.2	Digitale Welt . . . . .	6
<b>4</b>	<b>Ziel</b>	<b>7</b>
4.1	Szenario . . . . .	7
4.2	Kernpunkte . . . . .	8
<b>5</b>	<b>Konzept</b>	<b>8</b>
5.1	Informationsgewinnung . . . . .	8
5.2	Informationsaufbereitung . . . . .	10
5.3	Validierung . . . . .	12
<b>6</b>	<b>Zusammenfassung und Risiken</b>	<b>12</b>
<b>A</b>	<b>Anhang</b>	<b>13</b>
A.1	RDF-Beispiel . . . . .	13
	<b>Literatur</b>	<b>14</b>

## 1 Einleitung

Diese Ausarbeitung handelt von der automatischen Erstellung eines Benutzerprofils mit dem Fokus auf Social Navigation. Sie soll einen Ausblick auf eine mögliche Masterarbeit in diesem Gebiet geben, die Vorgehensweise darstellen und die damit verbundenen Risiken aufzeigen. Zu Beginn wird erläutert, was der Begriff *Social Navigation* bedeutet. Anschließend werden in Kapitel 3 zwischenmenschliche Strukturen und Personen näher betrachtet, da diese einen direkten Einfluss auf Social Navigation haben. Hier wird zwischen „Realer Welt“ und „Digitaler Welt“ unterschieden.

Im Kapitel 4 wird das Ziel der Arbeit anhand eines Szenarios genauer dargestellt. Das Kapitel 5 beschreibt eine mögliche Vorgehensweise zur Erreichung des im vorherigen Kapitel dargestellten Ziels. Die Vorgehensweise wird durch die Unterkapitel 5.1 bis 5.3 von der Informationsgewinnung, Informationsaufbereitung bis zur Validierung konkretisiert.

Abschließend gibt Kapitel 6 eine Zusammenfassung des Projekts und stellt die Risiken dar.

## 2 Grundlagen

In der AW2-Ausarbeitung von Schmidt (2006) wurde bereits ausführlich auf Social Navigation eingegangen. Daher befasst sich Kapitel 2.1 dieser Ausarbeitung lediglich mit dem Begriff und den Vorteilen dieser Navigationsart. Anschließend wird im Kapitel 2.2 erläutert, welche Rolle ein Benutzerprofil in diesem Zusammenhang hat.

### 2.1 Social Navigation

„Unter Social Navigation versteht man Konzepte, in denen sich Benutzer bei der Navigation am Verhalten und Hinweisen anderer Nutzer orientieren.“ Dourish und Chalmers (1998)/Dieberger u. a. (2000)

Eine Person hat also die Möglichkeit, anderen Personen Informationen zu hinterlassen und auf Informationen von anderen Nutzern zuzugreifen. Durch die Verwendung dieser Informationen kann auf das Verhalten bzw. die Entscheidung einer Person Einfluss genommen werden.<sup>1</sup>

Die Vorteile von Social Navigation sind, dass der Nutzer nicht mehr nur auf sich allein gestellt ist, sondern zusätzliche Informationen zur Entscheidungsunterstützung hat. Zudem haben Dourish & Chalmers herausgefunden, dass das Wohlbefinden durch diese Navigationsart steigt.<sup>2</sup>

### 2.2 Benutzerprofil

Für die effektive Nutzung von Social Navigation ist ein Benutzerprofil notwendig. Es wird unter anderem benötigt, um:

- persönliche Daten abzuspeichern
- zurückverfolgbare Spuren und Hinweise hinterlassen zu können
- mit anderen Benutzern in Kontakt zu treten
- individuelle Vorschläge und Hinweise erhalten zu können

Damit ein Benutzer relevante und individuelle Vorschläge und Hinweise erhalten kann, muss er eindeutig wiedererkennbar sein, d.h. es müssen persönliche Daten wie z.B. Name, Anschrift

---

<sup>1</sup>Vgl. Schmidt (2006)

<sup>2</sup>Vgl. Dourish und Chalmers (1998)

und/oder E-Mail-Adresse über ihn bekannt sein. Für den Benutzer sind alle Hinweise interessant, die für ihn und sein Vorhaben von Bedeutung sind. Dafür muss z.B. herausgefunden werden, welche Vorlieben der Benutzer hat.

Wenn kein Profil und somit keine Informationen über den Benutzer vorhanden sind, kann kein individueller Vorschlag gemacht werden, und ein Benutzer könnte nur verallgemeinerte Vorschläge erhalten.<sup>3</sup>

## 3 Zwischenmenschliche Strukturen und Personen

Im diesem Kapitel geht es um zwischenmenschliche Strukturen und Personen. Das Kapitel 3.1 zeigt den Blickwinkel der „Realen Welt“ aus Sicht der Soziologie. Hier werden verschiedene zwischenmenschliche Strukturen erläutert und die Merkmale von Personen betrachtet. Im Kapitel 3.2 wird der aktuelle Stand der Forschung vorgestellt. Zwei Projekte zeigen beispielhaft, inwieweit es möglich ist, die im vorhergehenden Kapitel beschriebenen Strukturen und Merkmale, in maschinenlesbarer Form abzubilden.

### 3.1 Reale Welt

In diesem Kapitel werden zunächst die zwischenmenschlichen Strukturen und anschließend die Merkmale von Personen aus soziologischer Sicht dargestellt.

#### Zwischenmenschliche Strukturen

Um aufzuzeigen, welche Arten von zwischenmenschlichen Strukturen existieren bzw. für diese Ausarbeitung von Interesse sind, wird hier zunächst der Ansatz der Dreiertypologie von Luhmann und anschließend die Erweiterung dieser durch den Gruppenbegriff von Neidhard vorgestellt.

Die Dreiertypologie sozialer Systeme ist Mitte der 70er Jahre entstanden und umfasst Interaktion, Organisation und Gesellschaft. Interaktionssysteme sind einmalige Zusammenschlüsse, die zusammenbrechen, wenn die Anwesenden das Terrain verlassen. Die Organisationen beruhen auf formaler Mitgliedschaft (z.B. Vereine) und die Gesellschaft wird als entpersonalisiertes Funktionssystem (z.B. Wirtschaft, Recht, Wissenschaft) angesehen.<sup>4</sup>

Da sich in dieser Typologie keine informalen Zusammenschlüsse wie z.B. Freundschaften unterbringen lassen, wurde die Dreiertypologie Ende der 70er Jahre von Willke, Neidhardt und Tyrell um den Gruppenbegriff erweitert.<sup>5</sup>

---

<sup>3</sup>Vgl. Schmidt (2006)

<sup>4</sup>Vgl. Luhmann (1991)

<sup>5</sup>Vgl. Neidhardt (1999)

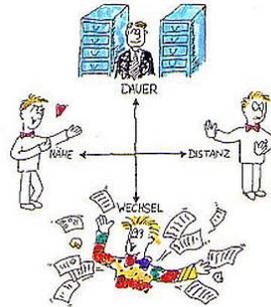


Abbildung 1: Riemann-Thomann-Modell

„Eine Gruppe ist ein soziales System, dessen Sinnzusammenhang durch unmittelbare und diffuse Mitgliederbeziehungen sowie durch relative Dauerhaftigkeit bestimmt ist.“ [Neidhard \(1999\)](#)

Diese Definition grenzt sich durch die relative Dauerhaftigkeit zum Interaktionssystem und durch die Unmittelbarkeit und Diffusität zu Organisationen ab.

### Personen

Dieser Abschnitt erläutert, was unter dem Begriff *Person* zu verstehen ist, bzw. welche für diese Ausarbeitung interessanten Merkmale eine Person besitzt.

Zu den Merkmalen einer Person gehören unter anderem sowohl demografische Daten als auch Persönlichkeitsmerkmale.

Die demografischen Daten umfassen die objektiv zu beobachtenden Merkmale. Hierzu zählen: Alter, Geschlecht, Wohnort, Einkommen etc. Diese Merkmale sind relativ einfach zu bestimmen und werden deshalb nicht genauer betrachtet.

Um die Persönlichkeitsmerkmale von Menschen zu bestimmen gibt es in der Literatur verschiedene Modelle und Theorien. Nachfolgend werden die drei Modelle „Riemann-Thomann-Modell“, „Big Five“ und „Myers-Briggs-Typindikator“ (MBTI) kurz vorgestellt.

- „Riemann-Thomann-Modell“  
Dieses Modell soll helfen, die menschlichen Unterschiede und ihre Auswirkungen auf Kommunikation und Beziehungen zu verstehen. Es ist ein Modell für die Persönlichkeits-, Beziehungs- und Entwicklungslehre und teilt die Bedürfnisse einer Person in vier verschiedenen Grundausrichtungen ein: Nähe, Distanz, Dauer und Wechsel (siehe Abbildung 1). Je nach Ausprägung sind entsprechende Bedürfnisse und „Lebensphilosophien“ vorherrschend, die sich im zwischenmenschlichen Verhalten widerspiegeln.<sup>6</sup>

<sup>6</sup>Vgl. [von Thun \(1998\)](#)

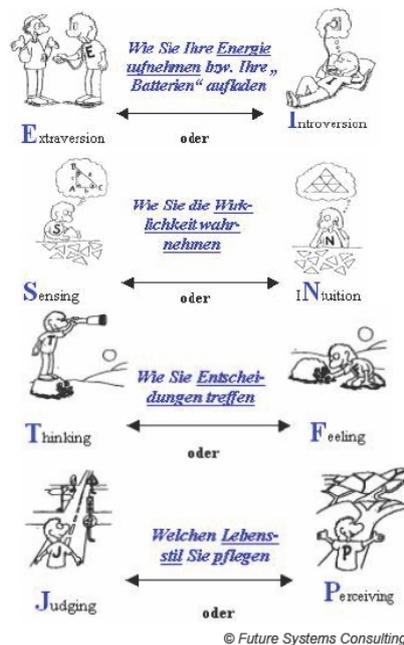


Abbildung 2: MBTI-Modell

- „Big Five“

Beim „Big Five“ bzw. „NEO-Fünf-Faktoren-Inventar“ handelt es sich um ein Modell, das aus fünf empirisch begründeten Persönlichkeitsfaktoren besteht. Auf Basis von Adjektivlisten und durch mehrfache Faktorisierung wurden fünf stabile, unabhängige Faktoren gefunden. Extraversion, Verträglichkeit, Gewissenhaftigkeit, Neurotizismus und Offenheit für Erfahrungen. Durch einen entsprechenden Fragebogen soll somit eine Erfassung der Persönlichkeit möglich sein.<sup>7</sup>

- „Myers-Briggs-Typindikator“ (MBTI)

Der MBTI dient zur Einschätzung der Persönlichkeit und stuft Menschen, bzw. deren Persönlichkeit in MBTI-Typen ein. Diese Typen bestehen aus vier Grundpräferenzen: Beziehungen zu anderen (**I**ntroversion/**E**xtraversion), Denkweise/Wahrnehmung (**S**ensing/**I**ntuition), Entscheidungen (**F**eeling/**T**hinking) und Lebensstil (**J**udging/**P**erceiving). (siehe Abbildung 2). Somit kann es 16 verschiedene MBTI-Typen (Vierertupel) geben. Das Tupel für eine Person, die extravertiert, praktisch veranlagt, objektiv und organisiert ist wäre z.B. ESTJ.<sup>8</sup>

<sup>7</sup>Vgl. Brokenau und Ostendorf (1993)

<sup>8</sup>Vgl. Myers (1985)

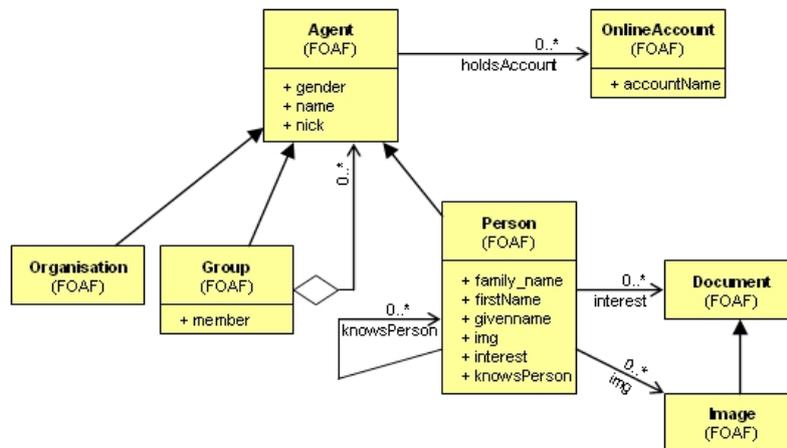


Abbildung 3: FOAF-Kurzübersicht (unvollständig) aus Domagalski (2006)

## 3.2 Digitale Welt

In diesem Kapitel wird der aktuelle Stand der Forschung betrachtet. Es wird anhand von zwei Projekten (FOAF und SOUPA) erläutert, inwieweit es möglich ist, die im vorhergehenden Kapitel beschriebenen Strukturen und Merkmale, in maschinenlesbarer Form abzubilden.

Die beiden Projekte FOAF und SOUPA sind sehr umfangreich. Daher wird in dieser Ausarbeitung nur ein grober Überblick über die Ontologien gegeben und die Teile, die für den weiteren Teil von Interesse sind, hervorgehoben.<sup>9</sup>

### FOAF (Friend of a Friend)<sup>10</sup>

Das Ziel von FOAF ist die Beschreibung von Personen und deren sozialen Strukturen. Mit FOAF ist es möglich die Beziehung von Personen, Personengruppen, Organisationen und Unternehmen maschinenlesbar abzubilden. Das zentrale Element ist die Beschreibung von Personen, die als Verbindung zwischen Personen, Dokumenten und Bildern fungieren (siehe Abbildung 3). FOAF verwendet die Beschreibungssprache RDF (Resource Description Framework), um Informationen zu speichern und auszutauschen. Eine solche RDF-Datei wird direkt bei der jeweiligen Ressource abgelegt und kann leicht gefunden werden.

Ein Beispiel einer Personen-RDF-Datei mit zugehöriger Gruppen-RDF-Datei ist im Anhang A.1 zu finden.

### SOUPA (Standard Ontology for Ubiquitous and Pervasive Applications)<sup>11</sup>

<sup>9</sup>Für weiterführende Informationen wird auf die Fachliteratur verwiesen.

<sup>10</sup>Vgl. FOAF (02.2007)

<sup>11</sup>Vgl. Chen u. a. (2004)

Die Entwicklung von pervasiven Anwendungen wird von SOUPA unterstützt. Es ist möglich, zusätzlich zu der Modellierung von Personen auch Zeit, Ort und Ereignisse zu modellieren.<sup>12</sup>

SOUPA setzt auf die Verwendung von bereits vorhandenen Ontologien. Diese werden durch OWL-(Web Ontology Language) Strukturen auf SOUPA-Konstrukte abgebildet. Unter anderem werden folgende Ontologien verwendet:

- FOAF für die Modellierung von persönlichen Daten und Beziehungen
- DAML-Time für die Beschreibung der Zeit
- OpenCyc für die Beschreibung von Orten

### **Zusammenfassung**

Die Darstellung der beiden Projekte zeigt, dass es schon bestehende Konzepte gibt, um personenbezogene Informationen in maschinenlesbarer Form auszudrücken. Durch FOAF ist es möglich, persönliche Daten, Organisationen, Gruppen und Beziehungen abzubilden. SOUPA ermöglicht zudem die Unterstützung für mobile Anwendungen.

## **4 Ziel**

Im diesem Kapitel wird das Ziel der Ausarbeitung anhand eines Szenarios konkretisiert. Anschließend werden die Kernpunkte der Arbeit herausgestellt.

### **4.1 Szenario**

Ein Benutzer plant die CeBIT zu besuchen und „surft“ mit seinem Browser auf den CeBIT-Seiten. Anhand seines Verhaltens wird automatisch ein Profil für ihn erstellt, das heruntergeladen und auf ein mobiles Gerät transferiert werden kann.

Durch dieses Profil erhält eine Anwendung folgende Möglichkeiten:

1. Auf der CeBIT-Karte kann angezeigt werden, welche Stände für den Benutzer interessant sind bzw. sein könnten.
2. Sobald sich der Benutzer einer anderen Person nähert, die ein ähnliches Interessengebiet hat, wird er über einen potentiellen Gesprächspartner informiert, sofern beide dies möchten.
3. Es kann angezeigt werden, an welchem Ort eine, für den Benutzer interessante Besprechung oder Diskussion stattfindet.

<sup>12</sup>Es sind weitere Modellierungsarten vorhanden. Siehe hierzu [Chen u. a. \(2004\)](#)

## 4.2 Kernpunkte

Ziel dieser Arbeit ist nicht eine Anwendungssoftware für die CeBIT zu entwickeln. Vielmehr wird für dieses Szenario davon ausgegangen, dass eine Anwendung vorhanden ist, die das Profil nutzen kann.

Aus dem Szenario ergeben sich somit folgende Kernpunkte für die automatische Erstellung eines Benutzerprofils:

- Es müssen Interessensgebiete erkannt und gewichtet werden. Dies bezieht sich auf alle drei der im Szenario genannten Punkte. Um diese Gebiete zu erkennen, müssen geeignete Daten auf Grundlage der CeBIT-Seite gesammelt und ausgewertet werden.
- Es soll überprüft werden, ob Persönlichkeitsmerkmale aus dem Verhalten einer Person ableitbar sind. Die Persönlichkeitsmerkmale sind deshalb von Interesse, da z.B. das Merkmal der Extaversion dafür sprechen würde, dass man potentiell daran Interesse hat mit anderen Menschen in Kontakt zu treten (siehe Punkt 2 des Szenarios).
- Es soll überprüft werden, ob Gruppenbeziehungen automatisch erkannt werden können. Dies korreliert mit Punkt 3 des Szenarios. Ein Zusammentreffen von mehreren Personen einer bestimmten Gruppe kann ein Indiz für eine interessante Besprechung sein, wenn der Benutzer dieser Gruppe angehört.

## 5 Konzept

In diesem Kapitel wird beschrieben, wie ein Benutzerprofil automatisch erstellt werden kann, das den Anforderungen, die in Kapitel 4 dargestellt worden sind, gerecht wird. Das Kapitel 5.1 beschreibt die Möglichkeiten der Informationsgewinnung. In Kapitel 5.2 werden Möglichkeiten zur Benutzerprofilerstellung und Auswertungsmethoden aufgezeigt. Das Kapitel 5.3 stellt dar, wie das automatisch erzeugte Benutzerprofil validiert werden kann.

### 5.1 Informationsgewinnung

In diesem Kapitel wird beschrieben, wie die Informationen gewonnen werden können, die für die Erstellung des Profils bzw. als Grundlage für die Auswertung dienen.

#### Allgemein

Wie im Szenario des Kapitels 4.1 werden die Daten auf Basis der CeBIT-Webseite gesammelt. Dies hat den Vorteil, dass der Benutzer schon Interesse bekundet hat und das Interessensgebiet des Benutzers eingegrenzt worden ist. Der begrenzte Suchraum hat den Vorteil, dass

die relevanten Begriffe bereits von der CeBIT-Seite vorgegeben und damit definiert sind. Die Schlüsselwörter, die zum Einordnen und Gruppieren der Interessen verwendet werden, sind also bekannt und für das Szenario ausreichend detailliert.

Ein Nachteil des eingeschränkten Suchraums ist, dass evtl. nicht alle Daten verfügbar sind, um alle Ziele des Szenarios abzudecken. Um Persönlichkeitsmerkmale erkennen zu können ist eine Ausweitung des Suchraums notwendig. Da der Fokus in diesem Szenario auf der Extraversion liegt, wird der Suchraum um die E-Mails des Benutzers erweitert. Diese clientseitige Erweiterung kann durch ein Plugin für den Browser realisiert werden.

### **Serverseitige Informationsgewinnung**

Ohne dem Benutzer Fragen zu stellen, können folgende Informationen über ihn gewonnen werden:

- HTTP Request:
  - User Agent String
  - IP-Adresse
  - Referrer-Information
- Javascript/Client Side:
  - Display-Information
  - Cookie-Information
- Click-Stream:
  - angesehener Inhalt
  - Reihenfolge des angesehenen Inhalts (Pfad durch Seite)
  - ausgeführte Suchanfragen
  - Zeitinformationen

Der User Agent String gibt Auskunft über den verwendeten Browser und das jeweilige Betriebssystem, IP-Adresse über den Aufenthaltsort und den Provider des Benutzers. Die Benutzeridentifikation kann über Cookies vorgenommen werden.

Die Click-Stream-Informationen sind für Benutzerprofilierung am interessantesten, da sie die detailliertesten Informationen beinhalten. Suchanfragen des Benutzers sowie Schlüsselwörter können direkt ausgelesen werden. Zudem erlauben die Zeitinformationen (Datum/Uhrzeit und Zeit zwischen den Klicks) und die angesehenen Inhalte direkte Rückschlüsse auf die Verweildauer der jeweiligen Seite.

Der Vorteil bei der serverseitigen Informationsgewinnung ist, dass es bereits mehrere Produkte auf dem Markt gibt, die für dieses Vorgehen geeignet und nutzbar sind. Beispiele für

diese Programme sind: Log File Analyzer<sup>13</sup>, Web Suxess<sup>14</sup> und Analog<sup>15</sup>

Der Nachteil liegt bei der Benutzeridentifikation, da für diese eine Registrierung notwendig ist. Ein Cookie reicht hierfür nicht aus, da dies gelöscht und der Benutzer nicht mehr identifiziert werden kann. Ein weiterer Nachteil bei der rein serverseitigen Gewinnung ist, dass keine Informationen über Kontakte und Persönlichkeitsmerkmale des Benutzers zugänglich sind.

### **Clientseitige Informationsgewinnung**

Mit clientseitiger Informationsgewinnung ist hier das oben angesprochene Plugin für den Browser gemeint. Da das Plugin eine eigenständige Software ist, die in den Browser eingebettet wird, kann diese sowohl auf Browserfunktionen als auch auf externe Programme zugreifen.

Der Vorteil bei dieser Vorgehensweise ist, dass auf mehr Informationen zugegriffen werden kann. Das Programm kann z.B. auf den E-Mail-Client des Benutzers zugreifen und somit an E-Mails und Kontakte gelangen. Zudem ist eine detaillierte Aussage über die Aufmerksamkeit des Benutzers möglich. Zusätzlich zu der „Klickfrequenz“ ist z.B. die Mausbewegung abfragbar und es kann ermittelt werden, ob sich der Browser noch im Vordergrund befindet oder ob der Benutzer seine Aufmerksamkeit einer anderen Anwendung schenkt.

Der Nachteil liegt bei der Akzeptanz des Benutzers. Er muss dem Plugin gestatten auf sensible bzw. persönliche Daten zuzugreifen.

## **5.2 Informationsaufbereitung**

In diesem Kapitel wird beschrieben, welche Möglichkeiten vorhanden sind, um Benutzerprofile zu erstellen und welche Auswertungsmethode für diese Ausarbeitung geeignet ist.

### **Methoden zur Benutzerprofilerstellung**

Nach [Kuflik und Shoval \(2000\)](#) existieren sechs Methoden um Benutzerprofile zu erstellen.

- „User-Created Profile“  
Der Benutzer spezifiziert seine Interessensgebiete anhand einer Liste von (evtl. gewichteten) Begriffen selbst.
- „System-Created Profile by Automatic Indexing“  
Eine Menge von relevanten Daten wird von einer Software analysiert, um die häufigsten Begriffe zu identifizieren. Diese Begriffe werden gewichtet und das Profil wird entsprechend angereichert.
- „System- plus User-Created Profile“  
Eine Kombination der beiden vorangehenden Methoden. Zuerst wird ein initiales, auto-

---

<sup>13</sup>Vgl. [LogFileAnalyzer \(02.2007\)](#)

<sup>14</sup>Vgl. [WebSuxess \(02.2007\)](#)

<sup>15</sup>Vgl. [Analog \(02.2007\)](#)

matisch generiertes Profil erstellt, das anschließend vom Benutzer durch Hinzufügen, Ändern und Gewichten überarbeitet wird.

- „System-Created Profile based on Learning by Artificial Neural-Network (ANN)“  
Aufgrund einer großen Datenmenge wird das Neuronale Netz mit relevanten Daten angereichert. Anhand der Benutzereingaben wird die Relevanz dieser Daten bestimmt und geht in das zukünftige Profil ein.
- „User-Profile Inherited from a User Stereotype“  
Diese Methode bezieht sich auf vordefinierte Benutzer-Stereotypen, also virtuelle Benutzer, die bestimmte Eigenschaften besitzen. Ein Benutzer wird dem Stereotypen zugeordnet, der ihm am ähnlichsten ist.
- „Rule-based Filtering“  
Das Benutzerprofil wird durch Regeln auf Grundlage von vordefinierte Fragen oder Benutzer-Stereotypen erstellt.

Die Methoden „User-Created Profile“ und „System- plus User-Created Profile“ sind nicht geeignet, da das Profil automatisch und ohne Zutun des Benutzers erstellt werden soll. Der Benutzer soll also nicht aktiv Daten einpflegen. Die Methode „System-Created Profile based on Learning by Artificial Neural-Network (ANN)“ wird nicht verwendet, da das Neuronale Netz erst mit vielen Daten angelernt werden muss. Die Methoden „User-Profile Inherited from a User Stereotype“ und „Rule-based Filtering“ finden ebenfalls keine Verwendung, da bei ihnen das Benutzerprofil eher pauschalisiert als personalisiert wird.

Für diese Ausarbeitung wird die Methode „System-Created Profile by Automatic Indexing“ verwendet, da diese aufgrund der vorliegenden Daten und der Definition des automatischen Generierens (ohne Aktivität des Benutzers) am geeignetsten erscheint.

### **Auswertungsmethode**

Die Benutzerprofilerstellung durch „System-Created Profile by Automatic Indexing“ entspricht dem Web-Mining bzw. Data-Mining, da bei diesen der gleiche Ansatz im Vordergrund steht. Es werden unbekannte Zusammenhänge, Muster und Trends aus einem Datenbestand ermittelt. Somit können Regelmäßigkeiten, Auffälligkeiten oder komplexe Zusammenhänge gefunden werden, die für das Erstellen eines individuellen Profils notwendig sind.

Diese Methode zu benutzen hat zudem den Vorteil, dass Data-Mining eine verbreitete und anerkannte Technik ist und Libraries für unterschiedlichste Programmiersprachen verfügbar sind. Beispiele für einsetzbare Libraries sind: XELOPES<sup>16</sup> (Java, C++ und C#) und Weka3<sup>17</sup> (Java).

---

<sup>16</sup>Vgl. [XELOPES \(02.2007\)](#)

<sup>17</sup>Vgl. [Weka3 \(02.2007\)](#)

### 5.3 Validierung

Dieses Kapitel zeigt auf, mit welchen Methoden überprüft werden kann, ob das automatisch erstellte Benutzerprofil korrekt ist. Um diesen durchführen zu können, werden mehrere Probanden benötigt, die ihren Interessen auf den CeBIT-Seiten nachgehen und Zugriff auf ihre persönlichen Daten (Email-Client) gewähren.

#### Vergleich

Die Überprüfung auf inhaltliche Übereinstimmung von zwei Profilen kann mit Hilfe eines Vergleichs erfolgen. Die Probanden müssen einen anerkannten MBTI-Bewertungsbogen und einen Fragebogen zu ihren Interessen ausfüllen.

Die Ergebnisse der Fragebögen werden mit dem automatisch, aus dem Verhalten auf den Seiten und den Daten des Email-Clients, erstellen Benutzerprofil verglichen. Durch eine geeignet große Menge von Testpersonen können qualitativ hochwertige Aussagen über die Korrektheit des Systems getroffen werden.

#### Befragung

Eine weitere Möglichkeit ist, die Probanden das automatisch erzeugte Profil bewerten zu lassen, also die einzelnen Merkmale des Profils mit einer eigenen Selbstbewertung zu vergleichen.

Durch diese Möglichkeit hätte man zwar eine direkte Rückmeldung, allerdings muss hierbei beachtet werden, dass Personen sich selbst nicht objektiv einstufen.

## 6 Zusammenfassung und Risiken

In dieser Ausarbeitung wurde ein Vorgehen vorgestellt, nach dem ein Benutzerprofil für Social Navigation automatisch erstellt und auf Funktionalität überprüft werden kann. Es wurden bestehende Projekte und Programme vorgestellt, mit denen eine funktionale Realisierung möglich ist. Zudem wurden Methoden zur Erstellung des Benutzerprofils aufgezeigt und bewertet. Abschließend wurden Möglichkeiten für eine Validierung des Konzeptes aufgezeigt.

Das Konzept ist nicht ohne Risiken. Ergebnis der Validierung kann sein, dass die automatische Gewinnung/Erzeugung zu einem „falschen“ oder zu einem „unzureichenden“ Profil führt. Dies würde einen erfolgreichen Einsatz verhindern, da dem Benutzer keine hilfreichen Informationen und somit kein Mehrwert geliefert werden kann.

Zudem könnte sich im Verlauf der Arbeit herausstellen, dass es auf Grundlage der Daten nicht möglich ist, Rückschlüsse auf Persönlichkeitsmerkmale oder Gruppenbeziehungen zu ziehen. Dies würde zwar den Funktionsumfang des Projektes einschränken, es aber nicht grundsätzlich gefährden. Ein Ergebnis in diese Richtung könnte durchaus für weitere Arbeiten von Interesse sein.

## A Anhang

### A.1 RDF-Beispiel

Beschreibung einer Person in RDF:

```
<rdf:RDF xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#
xmlns:foaf="http://xmlns.com/foaf/0.1/">
<!-- Informationen zu einer Person: -->
<foaf:Person rdf:nodeID="ich">
<foaf:name>Max Mustermann</foaf:name>
<foaf:mbox rdf:resource="mailto:max@mustermann.de"/>
<foaf:myersBriggs>ESTJ</foaf:myersBriggs>
...
<!-- Die Person ist in folgender Gruppe: -->
<foaf:Group>
<foaf:name>Mustergruppe</foaf:name>
<rdfs:seeAlso rdf:resource="http://example.com/mustergruppe.rdf" />
<foaf:member rdf:nodeID="ich"/>
</foaf:Group>
</foaf:Person>
```

Beschreibung der zugehörigen Gruppe in RDF:

```
<foaf:Group>
  <foaf:name>Mustergruppe</foaf:name>
  ...
  <foaf:member>
    <foaf:Person>
      <foaf:name>Max Mustermann</foaf:name>
    </foaf:Person>
  </foaf:member>
</foaf:Group>
```

## Literatur

- [Analog 02.2007] ANALOG: (02.2007). – URL <http://www.analog.cx>
- [Brokenau und Ostendorf 1993] BROKENAU, P. ; OSTENDORF, F.: NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae. In: *Hogrefe-Verlag* (1993)
- [Chen u. a. 2004] CHEN, Harry ; PERICH, Filip ; FININ, Tim ; JOSHI, Anupam: SOUPA: Standard Ontology for Ubiquitous and Pervasive Applications. (2004)
- [Dieberger u. a. 2000] DIEBERGER, A. ; DOURISH, P. ; HÖÖK, K. ; RESNICK, P. ; WEXELBLAT, A.: Social Navigation: techniques for building more usable systems *Interactions*. (2000)
- [Domagalski 2006] DOMAGALSKI, Roland: *Entwurf einer Semantic Web Ontologie für kontextsensitive, adaptive Anwendungen*, Diplomarbeit, 2006
- [Dourish und Chalmers 1998] DOURISH, P. ; CHALMERS, M.: Running Out of Space: Models of Information Navigation. (1998)
- [FOAF 02.2007] FOAF: (02.2007). – URL <http://xmlns.com/foaf/01>
- [Kuflik und Shoval 2000] KUFLIK, Tsvi ; SHOVAL, Peretz: Generation of User Profiles for Information Filtering - Research Agenda. In: *ACM-Press* (2000)
- [LogFileAnalyzer 02.2007] LOGFILEANALYZER: (02.2007). – URL <http://www.surfstats.com>
- [Luhmann 1991] LUHMANN, Niklas: *Interaktion, Organisation, Gesellschaft*. Westdeutscher Verlag, 1991 (Soziologische Aufklärung 2)
- [Myers 1985] MYERS, Isabel B.: *A Guide to the Development and Use of the Myers-Briggs Type Indicator*, 1985
- [Neidhard 1999] NEIDHARD, Friedhelm: *Innere Prozesse und Außenweltbedingungen sozialer Gruppen*. 3. Quelle & Meyer, 1999 (Einführung in die Gruppensoziologie)
- [Schmidt 2006] SCHMIDT, Thomas: Social Navigation - Möglichkeiten und Grenzen -. (2006)
- [von Thun 1998] THUN, Schulz von: *Miteinander reden 3. Das Innere Team und situationsgerechte Kommunikation*. Rowohlt Tb., 1998
- [WebSuxess 02.2007] WEBSUXESS: (02.2007). – URL <http://www.exody.net>
- [Weka3 02.2007] WEKA3: (02.2007). – URL <http://www.cs.waikato.ac.nz/ml/weka>
- [XELOPES 02.2007] XELOPES: (02.2007). – URL <http://www.prudsys.de/Software/Algorithmen/Xelopes>