



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Seminar

Lars Mählmann

Deliever who I mean continues ...

Lars Mählmann
Deliever who I mean continues ...

Seminar eingereicht im Rahmen der Diplomprüfung
im Studiengang Informatik
Studienrichtung Softwaretechnik
am Fachbereich Elektrotechnik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer :
Zweitgutachter : Prof. Dr. Kai von Luck

Abgegeben am 5. März 2008

Inhaltsverzeichnis

1	Einleitung	3
2	Szenario	4
3	Social Networking	5
3.1	Verwandte Arbeiten	6
4	Design	6
4.1	Frontend	7
4.2	Backend	9
4.3	Zusammenfassung	11
5	Ausblick	11
5.1	aktueller Stand	11
5.2	Problemstellungen	12
5.3	Lösungsansatz	12
5.4	Fazit	12
	Literatur	13

1 Einleitung

Das Auffinden der wichtigen und relevanten Informationen ist seit jeher ein schwieriges Thema. Konventionelle Ansätze, wie z.B. Datenbanken, "information retrieval systems" oder Internetsuchmaschinen, haben verschiedene Lösungen entwickelt und bis zu einem bestimmten Grad gelöst. Allerdings stellt sich immer wieder das Problem, daß die relevanten Informationen in einem Inter- oder Intranet nicht auffindbar sind und man auf das Wissen von Bekannten, Freunden oder, in Firmen, auf Mitarbeiter angewiesen ist (Yu und Singh (2002), ¹). Sind diese Sozialen Netzwerke nicht vorhanden, sind die gesuchten Informationen nicht verfügbar. Um die Problemstellung des Aufbaus von Sozialen Netzwerken zu lösen, gibt es verschiedene Ansätze, wie die Anwendungen Facebook, MySpace, oder Xing, die versuchen eine "Friend of a Friend" Struktur aufzubauen. Die Verbreitung von möglichen relevanten Informationen durch "Social Navigation" ist ein weiterer Ansatz. Es ermöglicht die Profitierung von Erfahrung anderer für den eigenen Zweck. Beispiele hierfür sind Amazon ² (".. Kunden die diesen Artikel gekauft haben, ...") oder das Austauschen von Bookmarks, wie bei del.icio.us

¹Clemens, Managing Enterprise Content: A Unified Content

²<http://amazon.com>

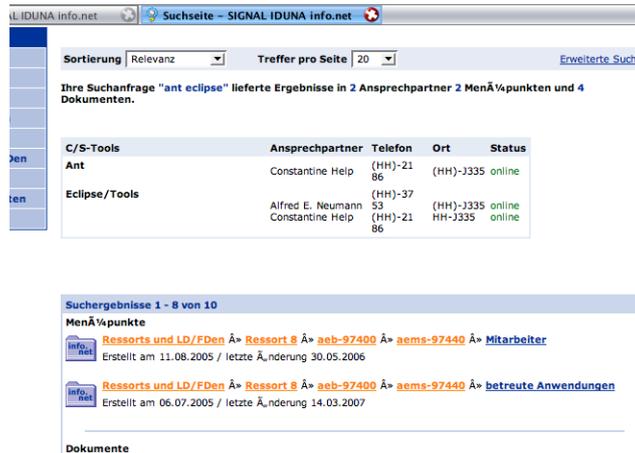


Abbildung 1: Abbildung einer möglichen Suchmaske

social bookmarking system³. Es profitieren andere Benutzer von den eigenen sortierten Bookmarks. Das Auffinden von Personen, die im Besitz von relevanten Informationen sind, soll in diesem Bericht für den Fall eines Intranets untersucht werden.

Diese Ausarbeitung stellt den aktuellen Stand vor. Es wird das Szenario, die Problemstellung 2 dargelegt und einzelne Konzepte 3.1 vorgestellt. Der aktuelle Stand wird dokumentiert und ein Ausblick auf die anstehende Arbeit einer möglichen Masterarbeit gegeben 4. Weiterhin werden die verschiedenen Risiken 5.2 betrachtet und ein Ausblick auf das Ziel dieser Arbeit gegeben 5.3.

2 Szenario

Im Folgenden wird eine kurze Beschreibung der Problemlage gegeben. Es wird dargestellt, wie die Anforderung aussieht.

Es soll in einem Intranet die Suche verbessert werden, indem eine Erweiterung zur Suche nach Personen implementiert wird. Die Implementierung der Suche hat verschiedene Schwierigkeiten. Sie soll aus dem Intranet aus für jeden Benutzer anwendbar sein und in die bestehende Suche eingebunden werden können. Diese Anwendung könnte wie Grafik 1 aussehen.

Daraus ergeben sich die folgenden Fragen:

- Welche Informationen sind relevant, verfügbar?

³<http://del.icio.us/about/>

- Woher kommen die Informationen?
- Wie können diese Daten aus dem System extrahiert werden?
- Wie erhält man aussagekräftige Informationen aus der Menge an Informationen?

Es sollen Benutzerprofile angelegt werden, die einen Mitarbeiter beschreiben, aber nicht den Datenschutz verletzen. In der Beschreibung können Informationen z.B. festgehalten werden, z. B. über die Firmenstruktur, in der sich der Mitarbeiter befindet, sowie dem sozialen bzw. technischen Kontext. Die Informationen sollen automatisiert erstellt und immer aktuell gehalten werden.

Ähnliche Problemstellung wurden in verschiedenen Artikeln, wie [Tang u. a. \(2007a\)](#), [Tang u. a. \(2007b\)](#), [Farrell u. a. \(2007\)](#) oder [Yu und Singh \(2002\)](#) beschrieben und weiterausgeführt. In den folgenden Kapiteln soll ein Lösungsansatz vorgestellt werden der sich teilweise in der Ausarbeitung befindet (siehe [Mählmann \(2008\)](#)) und als Ansatz für eine Masterarbeit gilt.

3 Social Networking

Es wurde in der Einleitung von einem "Social Network" gesprochen, ohne aber eine genauere Erklärung zu dem Begriff zu geben. Auf den Begriff soll im Folgenden kurz eingegangen und einige Beispiele für "Social Networking" gegeben werden. Es beschreiben [Jamali und Abolhassani \(2006\)](#) wie folgt:

Social network is a set of people connected by a set of social relationships such as friendship, co-working or information exchange. . . .

"Social Networking" beschreibt also die Beziehung zwischen einzelnen Personen oder Gruppen. Weiterhin wird geschrieben,

Social network analysis focuses on the analysis of patterns of relationships among people, organizations, states and such social entities.

Die Beziehungen werden dabei über Interessensschwerpunkte gebildet wie Arbeit, Freundschaft, etc. . . . Bei einer Untersuchung von "Social Networks" fanden [Li u. a. \(2007\)](#) über 30 verschiedene Relationen. Der Aufbau eines Netzwerkes kann daher sehr komplex werden und es kann passieren, daß es zuviele Beziehungen zwischen jedem einzelnen Mitarbeiter geben kann (siehe [Matsuo u. a. \(2006\)](#)).

3.1 Verwandte Arbeiten

Es gibt verschiedene Forschungsprojekte, die sich mit dem Aufbau eines "Social Networks" beschäftigt haben. Das Projekt Polyphonet [Matsuo u. a. \(2006\)](#) ist eine Anwendung, die aus verschiedenen verfügbaren Dokumenten (verschiedene Datenformate) ein "Social Network" an Hand von Autoren, etc. . . . aufbaut. Fringe (siehe [Farrell u. a. \(2007\)](#)) ist ein weiteres Projekt in dem versucht wird mit Hilfe von Tags bestimmten Thematiken einer Person zuzuordnen und somit Beziehungen aufzubauen zwischen den einzelnen Mitarbeitern. Das Dashboard versucht relevante Informationen zu einem Topic in einem extra Fenster darzustellen, wie z. B. bei einer eMail sämtliche Informationen zu dem Absender ([Friedman \(2003\)](#)). Die Abbildung ?? zeigt Beispiele für die Anwendungen.

Diese Anwendung stellt Beispiele für "Social Networking", die Daten syntaktisch untersuchen (wie [Matsuo u. a. \(2006\)](#)) oder auf eine semantische Einordnung des Benutzers durchs klassifizieren mit Tags (siehe [Farrell u. a. \(2007\)](#)) basieren. In dem vorgestellten Intranetszenario soll versucht werden, alle möglichen Informationen syntaktisch zu untersuchen und diese zusätzlich mit einer semantischen Bedeutung anzureichern. Im anschliessenden Kapitel soll ein mögliches Konzept vorgestellt werden.

4 Design

Wie sollte eine Architektur für eine Personensuche aussehen? In diesem Kapitel soll ein Konzept für diese Architektur vorgestellt und erklärt werden.

Das System besteht aus zwei Teilen, einem Backend und einem Frontend. In dem Backend werden die Daten aus dem Intranet extrahiert, in einem Datawarehouse abgespeichert und ausgewertet. Das Frontend ist die Suchmaschine mit der ein Benutzer Personen findet. Die Abbildung 2 zeigt wie die einzelnen Bestandteile zusammenhängen und miteinander verbunden sind.

Diese beiden Komponenten werden im Folgenden genauer erläutert. Es wird die Idee kurz vorgestellt und die Methoden bzw. Tools mit denen man diese umsetzt.

Das Frontend wird als erster Schritt erklärt und anschliessend über dessen aktuellen Stand berichtet. Im zweiten Teil des Abschnitts wird das Konzept des Backends dargestellt und ein Ausblick auf die bevorstehende Masterarbeit zu diesem Themenbereich gegeben.

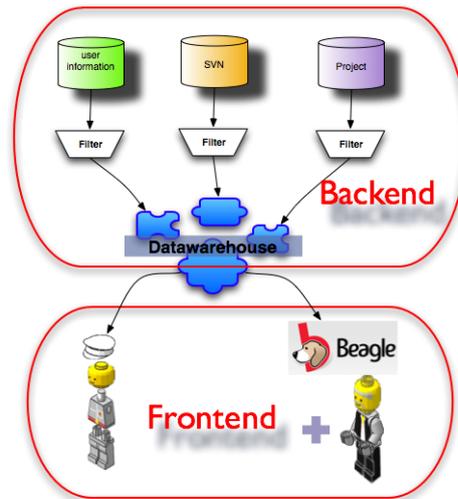


Abbildung 2: Abbildung einer möglichen Suchmaske

4.1 Frontend

Das Frontend stellt die Suchmaschine für die Personensuche. Die ausgewerteten Daten aus dem Datawarehouse müssen in eine Struktur gebracht werden, die einfach zu generieren ist und zusätzlich von einer Suchmaschine indiziert werden kann. Diese Suchmaschine stellt eine weitere Herausforderung dar. Ist es möglich eine bereits existierende Suchmaschine zu nutzen oder muss eine eigene Suche implementiert werden?

Die ausgewerteten Informationen können in einem Profil⁴ abgespeichert werden. Profile, werden häufig, in unterschiedlichen Kontexten (z. B. UML, Instant Messenger, Internetbrowser, etc , . . .) genutzt. Das Profil muss von einem Computer lesbar sein, d.h. ein Programm kann automatisiert diese Profile auslesen und interpretieren. Die Struktur dieses Profils muss komplex genug sein, um sämtliche Beschreibungen aufnehmen zu können. Zusätzlich muss das Profil erweiterbar sein. Ist ein Profil über eine Person angelegt, muss es möglich sein, die Information automatisch zu ändern oder weitere Attribute hinzuzufügen. Das bedeutet ein Profil sollte einem Standard erfüllen, um von Computerprogrammen interpretierbar zu sein. Zudem muss es erweiterbar und anpassbar sein, um flexibel auf Veränderungen eingehen zu können.

Die Anforderung der Maschinenlesbarkeit wird durch Semantic Web [Berners-Lee \(2008\)](#). Es muss möglich sein, eine Profilbeschreibung in Semantic Web Darstellung vorzunehmen. Eine Hilfestellung gibt ein Unterprojekt vom Resource Description Framework (RDF, [Ivan Her-](#)

⁴Configuration settings and other data associated with an individual user or with a group <http://en.wikipedia.org>

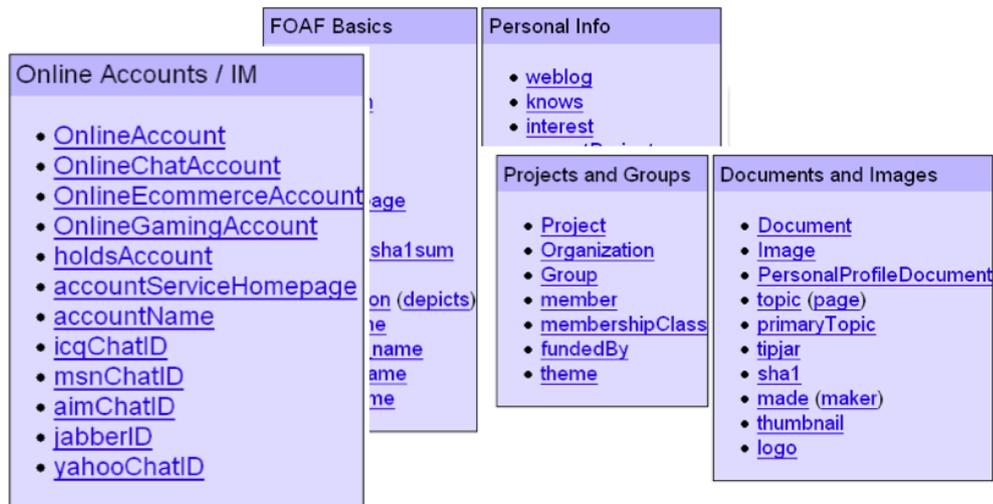


Abbildung 3: Abbildung eines FOAF Profiles

man und Brickley (2008)). Die Spezifikation von Friend of a Friend (FOAF Brickley und Miller (2008)⁵) beschreibt eine Profilbildung anhand von Semantic Web Spezifikationen. Es sind Attribute vorgegeben, die erweiterbar sind und zum Aufbau von "Social Networks" dienen sollen. Eine kleine Anzahl der verschiedenen Attribute ist in der Abbildung 3 zu sehen.

Im nächsten Schritt wird eine Suchmaschine benötigt, die ein solches Profil für eine Suche lesen, interpretieren und indizieren kann. Es gibt verschiedene "Query Languages", wie SPARQL Prud'hommeaux und Seaborne (2008) oder Suchmaschinen, die darauf aufbauen, wie z.B. Swoogle⁶. Eine weitere Möglichkeit wäre eine Desktopsuchmaschine so zu erweitern, daß diese diese Profile nicht als einfache Dateien ansieht, sondern deren Bedeutung versteht. Der Vorteil besteht darin, daß weiterhin die Standard Suche benutzt werden kann, es muss keine extra "Query Syntax" gelernt werden. Eine freiverfügbare Suchmaschine ist Beagle Love und Kubasik (2008). Sie basiert auf der Suchmaschine Lucene, die als OpenSource auf apache.org zur Verfügung gestellt wird.

Beagle ?? ist eine .Net Implementierung einer Desktopsuchmaschine, die durch zahlreiche Erweiterungen die Möglichkeit besitzt, verschiedene Dokumententypen unterschiedlich zu interpretieren und die Suche dadurch zu vereinfachen. Die Suchmaschine ist Komponentenbasiert aufgebaut und bietet eine Schnittstelle, um eigene Erweiterungen aufzunehmen. Beagle unterstützt XML Formate wie RDF und lässt sich somit einfach für eine Profilsuche auf FOAF Basis erweitern.

⁵<http://www.foaf-project.org/>

⁶http://swoogle.umbc.edu/index.php?option=com_swoogle_manual&manual=search_overview

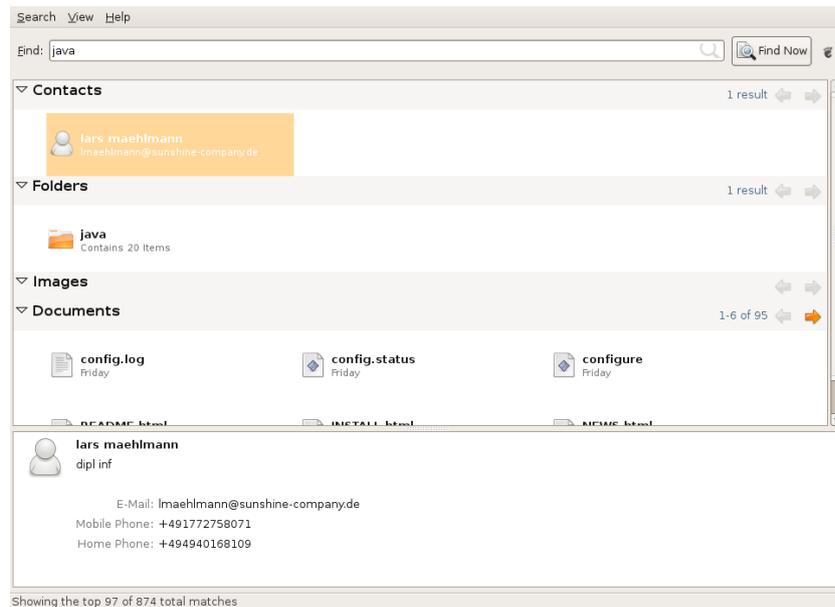


Abbildung 4: Das Suchergebniss nach dem Begriff Java

Der aktuelle Stand des Frontend zeigt, daß es möglich ist mit Beagle nach Profilen zu suchen, manuell ein Profil mit FOAF gebaut werden kann und dieses mit Beagle ausgelesen wird. Ein Beispiel dafür zeigt die Abbildung 4.

4.2 Backend

Im Backend sollen die Profile generiert werden, die zur Suche benutzt werden sollen. Die Informationen für ein Profil muss aus dem Intranet herausgefiltert werden. Ein Problem stellen die Daten dar, woher kann ich die Daten genau nehmen und welche Daten sind relevant. Kann man eine zeitliche Bewertung der Informationen vornehmen (z. B. eine Person hat vor 5 Jahren Wissen erworben und verfügt heute nicht mehr über das aktuelle benötigte Wissen.)? Das Intranet setzt sich aus mehreren Teilen zusammen. Es gibt Personalinformation (Adressen, Telefonnummern, etc . . .), Projektunterlagen, Repositories, Dokumente (z. B. Publikationen, Dokumentationen, Online Bücher, etc . . .), Web 2.0 Anwendungen (Blog, Wiki) und Firmeninterna wie Schulungsunterlagen und Informationen über Mitarbeiter Projekte, die nicht direkt abrufbar sind. In der Grafik 5 ist exemplarisch eine solche Struktur dargestellt (Clemens Grafik aus ...). Es gilt bei diesen Daten zu beachten, daß der Datenschutz in einer Firma nicht verletzt werden darf und daher einige Bereiche nicht verwendet werden dürfen.

Im ersten Schritt müssen die Daten aus dem System extrahiert werden. Es gibt, wie be-

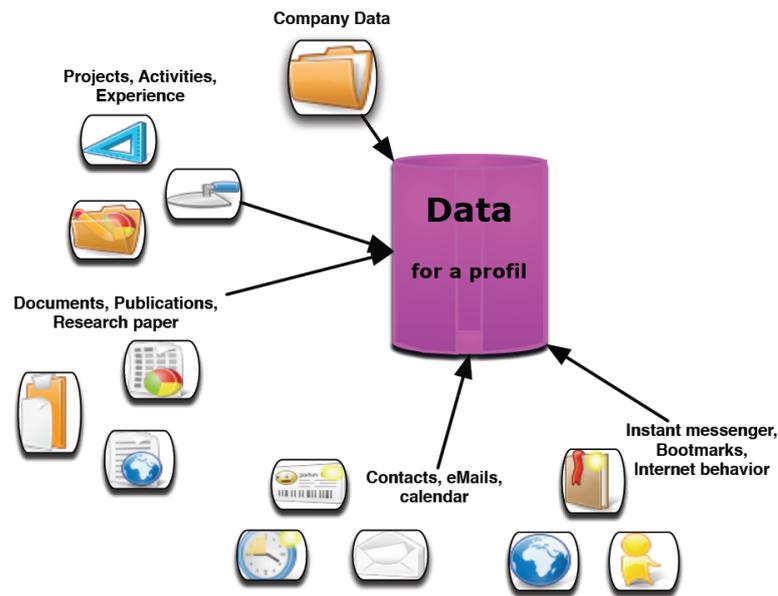


Abbildung 5: Darstellung der verschiedenen Bereiche eines Intranets

schrieben, verschiedene Datentypen, die unterschiedlich behandelt werden müssen. Die Daten können strukturiert vorhanden sein, wie z. B. ein LDAP Benutzerverzeichnis oder Datenbanken. Sie können aber auch vollständig unstrukturiert sein, wie ein SVN Repository, bei denen man aus dem Bezug zu einem Projekt keine Anhaltspunkte über die Daten hat. Eine letzte Möglichkeit ist die Informationsbeschaffung "zu Fuss". Es können Daten in der Art hinterlegt sein, daß kein direkter Zugang zu den Daten möglich ist und man sich selbst indirekt den Zugang beschaffen muss, wie z.B. bei Firmeninterna, die nur unter bestimmten Bedingungen herausgegeben werden. Um an diese Informationen zu gelangen, gibt es verschiedene Anwendungen die hier zum Einsatz kommen könnten. Als erstes könnte man eine Suchmaschine, wie Lucene [Cutting \(2008\)](#) anpassen, um an die Inhalte der Indizierung zu gelangen, womit relevante Begriffe herausgegriffen werden können. Mit diesen Begriffen kann man spezieller nach Informationen suchen. Eine weitere Möglichkeit sind Datenbankzugriffe mit denen der Zugriff auf Daten einfach ist, solange man weiß was man sucht. Es gibt hierbei weiterführende Literatur, wie z.B. Google hacks [Calishain und Dornfest \(2003\)](#) oder Arbeiten von [Matsuo u. a. \(2006\)](#) oder [Tang u. a. \(2007a\)](#), die sich mit der Gewinnung und Verarbeitung in diesem Bereich weiterführend beschäftigen.

In einem zweiten Schritt sollten die Daten mit Hilfe eines Datawarehouses aufbereitet werden. Es müssen die relevanten Informationen gefiltert werden, um dann später ein Profil zu generieren. Eine Schwierigkeit ist zu bestimmen, welche Informationen wichtig sind. Weiterhin reicht es nicht aus die Informationen zu finden, sondern diese müssen einer Person zugeordnet werden. Die Datenaufbereitung ist nicht vollständig geklärt, es gibt verschiedene

Ansätze und Verfahren die hierbei benutzt werden können, wie Association Analysis, Cluster Analysis, Text Mining oder wie von [Nair und Sarasamma \(2007\)](#) beschrieben Fuzzy Logic. Diese Verfahren sind erprobt und beschrieben. Zusätzlich gibt es dazu im Open Source Bereich verschiedene Programme, die zum Einsatz kommen können, wie "Pimiento" ([Adeva und Calvo \(2008\)](#)) , "Gate" ([hamish cunningham und diana maynard \(2008\)](#)) oder "Rapid-Miner" ([Rapid-I \(2008\)](#)). Ausserdem verändern sich die Daten über einen zeitlichen Raum, was bedeutet die Profile müssen in einem bestimmten Zeitraum überarbeitet werden und gegebenenfalls angepasst werden müssen. Mitarbeiter verlassen die Firma neue Mitarbeiter kommen hinzu. Der Schwerpunkt verschiebt sich, die Firmenstruktur ändert sich. Ein anderer Fall könnte sein, wenn man Schwerpunkte setzen möchte, in zeitlichen Abständen, mit denen sich verschiedene Mitarbeiter auseinander setzen und somit höher priorisiert werden bei ihrer Arbeit.

4.3 Zusammenfassung

In diesem Kapitel wurde ein Konzept zur Erstellung einer Personensuche vorgestellt. Einzelne Teile wurden genauer erläutert und die Probleme dargestellt, die in diesem Szenario zu beachten sind. Im nächsten Abschnitt soll ein Ausblick gegeben werden.

5 Ausblick

In der Zusammenfassung soll ein kurzer Iststand beschrieben werden. Die daraus entstandenen Probleme werden erklärt. Ein Ausblick auf die anstehende Masterarbeit in diesem Kontext gegeben, sowie ein Beschreibung des Zieles.

5.1 aktueller Stand

Bei der bisherigen Arbeit wurde das im vorherigen Kapitel 4 beschriebene Konzept als Basis zur Umsetzung benutzt. Der Schwerpunkt lag auf einer Implementierung der Personensuchfunktion und des Generieren eines Profils mit FOAF. Eine genauere Beschreibung ist im Projektbericht [Mählmann \(2008\)](#) nachzulesen. Das Ergebnis dieser Arbeit ist ein extra Filter für die Suchmaschine Beagle mit dem man die manuell erstellten FOAF Benutzerprofile indizieren kann und speziell nach ihnen sucht.

5.2 Problemstellungen

Wie schon im Kapitel 4.2 beschrieben gibt es verschiedene Probleme, die dazu führen können, daß es keine Personensuche geben kann. Das erste Problem besteht darin aus dem System relevante Daten zu extrahieren. Bei einer Datenbank ist es relativ einfach. Bei einem Projekt Verzeichnis ist es dagegen, aufgrund der Unstrukturiertheit und Unwissenheit über die Art der dort abgelegten Daten, schwierig. Das zweite Problem ist die Analyse der Daten mit Data Mining. Es besteht die Möglichkeit, daß die Ergebnisse zu abstrakt, zu detailliert sind oder es keine Ergebnisse gibt. Eine weitere Schwierigkeit besteht in der Zuordnung der Daten zu einer Person. Es ist möglich Personen zu finden und Information über eine Person zu finden, aber wie kann man diese Informationen verbinden und in einer Profilbeschreibung ablegen? Diese und verschiedene weitere Probleme sind in diversen Veröffentlichungen erläutert worden ⁷. Diese Schwierigkeiten sind bis jetzt nicht vollständig gelöst und erfordern weitreichende weitere Forschungsarbeit.

5.3 Lösungsansatz

Der Unterschied zu vorausgegangen Arbeiten besteht darin, daß ich versuche im Intranet, in eingeschränkter Form, eine Suche zu implementieren und in diesem nur ein kleinen Teil. Im Vergleich zu dem Internet ist das Intranet übersichtlicher. Es soll versucht werden zwei oder drei Filter für bestimmte Bereiche zu implementieren, wobei keine komplette Lösung gegeben werden soll. Die Informationen, die bearbeitet werden, sind bekannt und teilweise schon strukturiert abgreifbar. Dadurch sind die Informationen verifiziert und müssen nicht hinterfragt werden. Zusätzlich zu der Gewinnung von Informationen sollen diese ausgewertet und ein Profil auf Basis von FOAF automatisch generiert werden.

5.4 Fazit

Es wurde in dieser Ausarbeitung verschiedene Arbeiten, Konzepte und auch Teilergebnisse vorgestellt (z. B. das Frontend 4.1). Dies beschreiben den Aufbau und die Nutzung von Social Networking und beweisen, daß es möglich aus den bestehenden System ein Netzwerk aufzubauen. Es gibt aber verschiedene Schwierigkeiten, wie die Informationsgewinnung, die in weiteren Untersuchungen geklärt werden muss. Eine Masterarbeit in diesem Umfeld ist nicht ohne Risiko, wenn die Informationsgewinnung nicht gelingt. Aber auf Grund der bisher gewonnen Erfahrungen und den bereits erfolgten Arbeiten sollte es möglich sein in einem abgegrenzten Bereich eine Personensuche zu realisieren.

⁷Vgl. Kapitel 3.1

Literatur

- [Adeva und Calvo 2008] ADEVA, Juan José G. ; CALVO, Rafael: *Pimiento, Platform Independent Text Mining Engine Tool*. 2008. – <http://erabaki.ehu.es/jjga/pimiento/>
- [Berners-Lee 2008] BERNERS-LEE, Tim: *W3C, Semantic Web*. 2008. – <http://www.w3.org/2001/sw/pub>
- [Brickley und Miller 2008] BRICKLEY, Dan ; MILLER, Libby: *FOAF Vocabulary Specification 0.91*. 2008. – <http://xmlns.com/foaf/spec/>
- [Calishain und Dornfest 2003] CALISHAIN, Tara ; DORNFEST, Rael: *Google Hacks: 100 Industrial-Strength Tips and Tools*. Sebastopol, CA, USA : O'Reilly & Associates, Inc., 2003. – ISBN 0596004478
- [hamish cunningham und diana maynard 2008] CUNNINGHAM, valentin t. hamish ; MAYNARD diana: *Gate, General Architecture for Text Engineering*. 2008. – <http://gate.ac.uk/>
- [Cutting 2008] CUTTING, Doug: *Apache Lucene*. 2008. – <http://lucene.apache.org>
- [Farrell u. a. 2007] FARRELL, Stephen ; LAU, Tessa ; NUSSER, Stefan ; WILCOX, Eric ; MULLER, Michael: Socially augmenting employee profiles with people-tagging. In: *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*. New York, NY, USA : ACM, 2007, S. 91–100. – ISBN 978-1-59593-679-2
- [Friedman 2003] FRIEDMAN, Nat: Dashboard. In: *In Proceedings of the Linux Symposium Conference on the Linux kernel and major OS infrastructure and research projects*, 2003
- [Ivan Herman und Brickley 2008] IVAN HERMAN, Ralph S. ; BRICKLEY, Dan: *W3C, Resource Description Framework (RDF)*. 2008. – <http://www.w3.org/2001/RDF>
- [Jamali und Abolhassani 2006] JAMALI, Mohsen ; ABOLHASSANI, Hassan: Different Aspects of Social Network Analysis. In: *WI '06: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*. Washington, DC, USA : IEEE Computer Society, 2006, S. 66–72. – ISBN 0-7695-2747-7
- [Li u. a. 2007] LI, Rui ; BAO, Shenghua ; YU, Yong ; FEI, Ben ; SU, Zhong: Towards effective browsing of large scale social annotations. In: *WWW '07: Proceedings of the 16th international conference on World Wide Web*. New York, NY, USA : ACM, 2007, S. 943–952. – ISBN 978-1-59593-654-7

- [Love und Kubasik 2008] LOVE, Robert ; KUBASIK, Kevin: *Beagle Project*. 2008. – http://beagle-project.org/Main_Page
- [Mählmann 2008] MÄHLMANN, Lars: *Deliever who I mean*, Projektbericht, 2008
- [Matsumura u. a. 2005] MATSUMURA, Naohiro ; GOLDBERG, David E. ; LLORÀ, Xavier: Mining directed social network from message board. In: *WWW '05: Special interest tracks and posters of the 14th international conference on World Wide Web*. New York, NY, USA : ACM, 2005, S. 1092–1093. – ISBN 1-59593-051-5
- [Matsuo u. a. 2006] MATSUO, Yutaka ; MORI, Junichiro ; HAMASAKI, Masahiro ; ISHIDA, Keisuke ; NISHIMURA, Takuichi ; TAKEDA, Hideaki ; HASIDA, Koiti ; ISHIZUKA, Mitsuru: POLYPHONET: an advanced social network extraction system from the web. In: *WWW '06: Proceedings of the 15th international conference on World Wide Web*. New York, NY, USA : ACM, 2006, S. 397–406. – ISBN 1-59593-323-9
- [Nair und Sarasamma 2007] NAIR, Premchand S. ; SARASAMMA, Suseela T.: Data Mining Through Fuzzy Social Network Analysis. In: *North American Fuzzy Information Processing Society*, 2007, S. 251 – 255. – ISBN 1-4244-1214-5
- [Prud'hommeaux und Seaborne 2008] PRUD'HOMMEAUX, Eric ; SEABORNE, Andy: *SPARQL Query Language for RDF*. 2008. – <http://www.w3.org/TR/rdf-sparql-query/>
- [Rapid-I 2008] RAPID-I: *RapidMiner*. 2008. – <http://sourceforge.net/projects/yale>
- [Tang u. a. 2007a] TANG, John C. ; DREWS, Clemens ; SMITH, Mark ; WU, Fei ; SUE, Alison ; LAU, Tessa: Exploring patterns of social commonality among file directories at work. In: *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2007, S. 951–960. – ISBN 978-1-59593-593-9
- [Tang u. a. 2007b] TANG, John C. ; LIN, James ; PIERCE, Jeffrey ; WHITTAKER, Steve ; DREWS, Clemens: Recent shortcuts: using recent interactions to support shared activities. In: *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA : ACM, 2007, S. 1263–1272. – ISBN 978-1-59593-593-9
- [Yu und Singh 2002] YU, Bin ; SINGH, Munindar P.: *Searching Social Networks*. Raleigh, NC, USA : North Carolina State University at Raleigh, 2002. – Forschungsbericht